Computers and Education: Artificial Intelligence 10 (2026) 100587

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

Bias and representation in AI generated text-to-image in education: A systematic review

Lilach Alon [a,*], Dorit Hadar Shoval [b], Inbar Levkovich [a]

[a] Faculty of Education and Teaching, Tel-Hai, University of Kiryat Shmona in the Galilee, Kiryat Shmona, Israel
[b] Faculty of Social Sciences, Tel-Hai, University of Kiryat Shmona in the Galilee, Kiryat Shmona, Israel

ABSTRACT

AI generated text-to-image tools are rapidly entering educational settings, yet little is known about how their visual biases intersect with goals of equity, inclusion, and critical AI literacy. This systematic literature review maps and analyzes empirical studies that examine bias and representation in educational uses of AI generated text-to-image. Following PRISMA guidelines, we identified 31 peer reviewed studies published between 2023 and 2025 across K-12, higher education, and professional learning contexts. We used a six part analytic framework (gender; race, ethnicity, and socioeconomic status; culture and religion; age; body and (dis) ability; and content) to code how studies conceptualized and investigated bias. Across the corpus, biased representation was pervasive: images frequently centered white, male, Western, thin, and non disabled figures, while diversity related to age, body, and ability was largely overlooked. Most studies relied on image audits and qualitative methods, with few experimental or intervention based designs. This review is the first to synthesize how educational research conceptualizes, measures, and responds to bias in text-to-image tools' outputs. The findings reveal significant blind spots and highlight directions for research, design, and policy aimed at aligning generative AI with educational inclusion and critical AI literacy.

1. Introduction

Generative artificial intelligence (Gen-AI) is rapidly reshaping how images, text, and other media are produced and used in education. Alongside text based systems such as chatbots, AI generated text-to-image tools now allow teachers and students to create visual materials on demand by typing short prompts, with outputs that can be adapted to different subjects, age groups, and learning tasks (Bender, 2023; Chan & Hu, 2023). Early studies in art, design, and media education suggest that these tools can support creativity, experimentation, and project based learning, while work in medical and health professions education explores their potential to supplement traditional illustrations and clinical photographs (Bian et al., 2025; Cevik et al., 2025). At the same time, this rapid adoption has raised concerns about the accuracy, trustworthiness, and social implications of AI generated images in teaching and learning.

One central concern is bias in how AI generated images represent people, places, and knowledge. Research across journalism, medicine, and computer science shows that text-to-image models often reproduce narrow and unequal patterns of representation. They tend to over-represent white, male, Western, thin, and non disabled bodies in positions of authority, while marginalizing or stereotyping women, people of color, older adults, and other minoritized groups (Currie et al., 2024; Elsharif et al., 2025). These biases were not random. They arise from large-scale web image datasets on which text-to-image systems are trained and design choices that prioritize visual appeal and coherence over diversity and fairness.

Although there is growing technical and ethical work on bias in generative image models, the educational implications of these patterns are still poorly understood. Existing studies that use text-to-image tools in education are scattered across disciplines and mostly focus on feasibility, user perceptions, or pedagogical opportunities (Bender, 2023; Hu, 2025; Liao et al., 2025). Reviews of bias in text-to-image systems synthesize evidence from general purpose benchmarks and from domain specific applications such as journalism or medical imaging (Elsharif et al., 2025; Vázquez & Garrido-Merchán, 2024). In this context, general purpose benchmarks refer to standardized non educational evaluation datasets and test suites used in AI research, rather than evaluations situated in specific educational tasks, curricula, or classroom activities. Beyond representational fairness, emerging work raises concerns about a growing gap between the visual plausibility of AI generated images and their factual accuracy (Basty et al., 2025; Doumpioti & Huang, 2024). In educational settings, this becomes relevant when AI generated images are used as instructional exemplars without structured verification or critique, particularly in activities where learners rely on outputs to illustrate disciplinary concepts (Dakamsih & Alkayid, 2025; Ringvold et al., 2024; Romanik & Wegner, 2025).

* Corresponding author.
E-mail addresses: alonlil@telhai.ac.il (L. Alon), dorithad@telhai.ac.il (D. Hadar Shoval), levkovinb@telhai.ac.il (I. Levkovich).

https://doi.org/10.1016/j.caeai.2026.100587
Received 4 December 2025; Received in revised form 22 March 2026; Accepted 27 March 2026
Available online 27 March 2026
2666-920X/

To date, there is no systematic literature review that brings these strands together to examine how bias and representation are studied when AI generated images are used for teaching and learning. This systematic review addresses that gap. Our aim is to map and critically analyze how bias and representation are examined in educational uses of AI generated text-to-image tools. We focus on six key dimensions that are both prominent in the broader bias literature and central to educational equity: gender; racial, ethnic, and socioeconomic identity; cultural and religious affiliation; age; body and (dis)ability; and content related distortions. We synthesize how studies conceptualize and operationalize these forms of bias, which educational contexts and learner groups they address, and how educators and students perceive and respond to biased images. In addition, we analyze how patterns of bias vary across publication years, research designs, and AI tools, in order to understand how the evidence base is developing over time. Our analysis pays particular attention to which groups become hyper visible and which are nearly absent, and how these patterns intersect with educational equity.

By doing so, the review makes several contributions. Conceptually, it offers a structured lens for examining bias and representation in educational uses of text-to-image systems, grounded in both technical and educational research. Empirically, it consolidates a fragmented body of studies into a coherent map of current knowledge and blind spots, highlighting which bias dimensions, subject areas, and participant perspectives are well covered and which remain underexplored. Practically, it points to future directions for research, design, and policy. These include the need for more diverse and rigorous evaluation methods and for participatory approaches that center learners and teachers experiences with artificial intelligence generated images. They also include the development of guidelines that help educators adopt text-to-image tools in ways that support inclusion, foster critical AI literacy, and advance socially just teaching.

1.1. AI-generated text-to-image in education

AI generated text-to-image systems allow users to create images in response to short written descriptions. A user types a prompt, and the system produces one or more images that match this description (Bian et al., 2025). These models are trained on large collections of images and accompanying captions, which enables them to learn broad associations between words, visual elements, and styles (Vargas-Veleda et al., 2025). In widely used systems such as Stable Diffusion, images are generated by starting from random noise and gradually refining it until a picture that fits the prompt emerges, making it possible to produce many variations quickly and at low cost (Liao et al., 2025; Noel, 2024).

In education, text-to-image tools such as DALL·E, Midjourney, Stable Diffusion, and similar generators are now used alongside other generative AI systems. Studies describe their integration into visual arts, design, architecture, and other practice oriented fields to support creativity, visualization, and project based learning (Huh et al., 2025; Liao et al., 2025). In medical and health professions education, they are explored as a way to expand and adapt visual learning resources, while raising concerns about anatomical accuracy, diagnostic reliability, and realism (Cevik et al., 2025; Noel, 2024). Beyond content creation, text-to-image systems are introduced in media and communication courses as objects of inquiry that can prompt critical reflection on AI, creativity, and stereotypes (Bender, 2023; Hu, 2025).

This emerging work suggests that text-to-image tools are already integrated into diverse educational activities and settings. In educational settings, images often function as instructional exemplars that communicate who belongs in a domain and what roles, practices, and settings are treated as typical (Peterson et al., 2021; Reed et al., 2023; Vartiainen et al., 2025). This makes representational defaults and omissions in AI generated images pedagogically salient, even when studies do not measure downstream learning outcomes directly. They are used to spark creativity, visualize complex content, and rethink teaching strategies, while also opening up new spaces for critical discussion of AI. However, most studies focus on feasibility, user experience, or pedagogical opportunities and only touch briefly, if at all, on systematic questions of bias and representation.

Existing reviews of text-to-image systems synthesize technical and domain specific evidence, but they do not concentrate on educational contexts or on how learners and educators encounter and interpret biased images. At the same time, an emerging educational literature has begun to examine how biases in AI generated images appear in learning settings (Cooper & Tang, 2024; Thomson et al., 2025). However, this literature remains fragmented. Prior studies have tended to examine specific tools, prompts, or instructional contexts, rather than offering a broader account of how bias and representation are conceptualized and assessed across educational uses of text-to-image systems. As a result, there is still no systematic synthesis focused on how educational research has examined bias, representation, and their implications for the pedagogical use of these images.

1.2. Bias and representation in AI-generated text-to-image

Since text-to-image systems are trained on large collections of web images and captions, they inherit and may amplify the imbalances and stereotypes present in those datasets (Vargas-Veleda et al., 2025). Recent reviews show that gender, racial and ethnic bias are the most frequently documented, but that these often intersect with cultural, body related, age, and class based patterns (Elsharif et al., 2025; Vázquez & Garrido-Merchán, 2024). Across domains such as journalism, medicine, and design, studies report that text-to-image models tend to reproduce dominant beauty ideals, Western centric norms, and narrow occupational profiles, while underrepresenting or distorting minoritized groups (Currie et al., 2024; Gisselbaek et al., 2024).

Building on this literature and the patterns identified in our corpus, the present review focuses on six main dimensions of bias and representation that are especially salient for education: gender; racial, ethnic, and socio economic status; cultural and religious affiliation; age; body and (dis)ability; and content. In the following subsections, we use these six categories as an analytic lens for describing how bias and representation are conceptualized and examined in educational uses of text-to-image tools.

1.2.1. Bias 1: Gender

Gender bias appears in how text-to-image systems associate roles, activities, and visual styles with men and women. In medical imaging and anesthesiology, for example, DALL·E based images depict most professionals as male, with a disproportionately high share of light skinned men in senior roles, even in specialties where the real workforce is more gender balanced (Currie et al., 2024; Gisselbaek et al., 2024). Journalism and media analyses similarly find that women are more often portrayed in sexualized, aestheticized ways, while men are linked to authority, competence, and leadership (Vargas-Veleda et al., 2025). Systematic reviews of text-to-image bias identify gender as the most studied dimension and show consistent overrepresentation of men in high status professions and stereotypical feminization of care and service roles (Elsharif et al., 2025; Vázquez & Garrido-Merchán, 2024).

In this review, we use gender bias to refer to systematic patterns in how text-to-image outputs assign genders to learners, educators, and professionals, and how they visually code femininity and masculinity. We examine when and how the included studies identify gendered patterns in AI generated images, how they interpret these patterns in educational contexts, and to what extent they connect them to broader concerns about equity and participation.

1.2.2. Bias 2: Racial/ethnic/socio-economic status

Racial and ethnic biases are closely intertwined with gender and class. Text-to-image models often default to white protagonists when prompts do not specify race and produce narrow, sometimes caricatured depictions for prompts that reference particular regions or ethnic labels (Elsharif et al., 2025). In medical and health related image generation, both Currie et al. (2024) and Gisselbaek et al. (2024) report that people with darker skin tones are underrepresented or stereotypically portrayed, despite the diversity of real clinical workforces. Vázquez and Garrido-Merchán (2024) show how prompts such as "playing basketball", "a rich person", or "a poor person" trigger racialized associations, with some models repeatedly linking certain sports or socio economic positions to specific racial groups. These patterns reinforce longstanding associations between whiteness and professionalism or affluence, and between racialized groups and either poverty or physicality. In this review, we group racial, ethnic, and socio economic bias (SES) together to capture how text-to-image outputs visualize skin tone, facial features, clothing, language, and material surroundings. We code how studies document the racial and class composition of AI generated images, and whether they link these visual patterns to questions of educational access, opportunity, and belonging.

1.2.3. Bias 3: Culture/religious

Cultural and religious bias concerns how models depict places, practices, and beliefs. Elsharif et al. (2025) show that many text-to-image datasets and benchmarks remain heavily Western centric, and that prompts involving regions such as the Middle East or the Global South generate images that exaggerate particular items of dress, architecture, or religious symbolism. Their review stresses that cultural bias is often visible only when examined together with gender, race, and class, for example, when "a Middle Eastern teacher" is consistently portrayed as a man in specific clothing, while a neutral "teacher" prompt yields a white woman in a Western classroom. Work on AI assisted persona generation suggests that, without explicit cultural frameworks, generative systems risk defaulting to national or ethnic stereotypes, even when asked to represent diverse user groups (Avis et al., 2025). Thus, cultural bias reflects both the geographic skew of the training data and the tendency of models to lean on simplified cultural scripts when filling in contextual details. In this review, we treat cultural and religious bias as patterns in which text-to-image outputs depict language, dress, rituals, buildings, landscapes, and symbols associated with particular cultures or faiths. We examine whether educational studies explicitly assess the cultural situatedness of AI generated images and whether they use these outputs to foster or undermine culturally responsive teaching.

1.2.4. Bias 4: Age

Age bias has received less systematic attention than gender or race but emerges in several empirical studies. Analyses of professional images show that older adults are underrepresented compared with the age distribution of real workforces, and that when they do appear they are more likely to be shown as frail or passive rather than as experts or leaders (Gisselbaek et al., 2024). Elsharif et al. (2025) note that age often appears as a compounding factor in cultural and class based bias, for example, when "a poor person" or "a migrant" is depicted as either very young or very old in ways that align with familiar media narratives rather than with demographic reality. Such patterns subtly encode ageist assumptions about who is active, competent or technologically capable. In this review, we understand age bias as the underrepresentation, stereotyping, or invisibility of particular age groups in AI generated educational images. We code when studies comment on the age of depicted learners, teachers, and professionals, and how they connect these visual age profiles to ideas about expertise, authority, and the life course in education.

1.2.5. Bias 5: Body and (dis)ability

Body and (dis)ability bias refers to how different body sizes, shapes, and disabilities are represented, or omitted altogether. Vargas-Veleda et al. (2025) describe forms of "aesthetic violence" in AI generated images, where thin, white, conventionally attractive bodies dominate, and weight diversity is either erased or framed negatively. Their study of DALL·E based outputs finds systematic weight discrimination and the marginalization of non-normative bodies in positive or aspirational contexts. In medical imaging, Currie et al. (2024) report the absence of any visible disability markers among hundreds of generated medical professionals, highlighting how text-to-image outputs can normalize a non-disabled workforce. Broader media analyses point out that these patterns are not random: they reflect training data that privilege Eurocentric beauty standards and able bodied ideals and can have ethical implications by reinforcing stigma and limiting the range of bodies seen as professional, healthy, or desirable (Vargas-Veleda et al., 2025). In this review, we use body and (dis)ability bias to capture how AI generated educational images portray body size, shape, mobility, and disability. We record whether studies attend to the visibility or absence of disabled learners and professionals, and how they interpret body ideals in classroom or professional scenes.

1.2.6. Bias 6: Content

Content bias goes beyond who appears in an image to how situations, settings, and relationships are depicted. In journalism and visual culture, text-to-image systems have been shown to exaggerate certain visual tropes, dramatize scenes, or mix cultural symbols in ways that misrepresent social realities, even when demographic attributes look balanced (Vargas-Veleda et al., 2025). Domain specific studies in medicine reveal frequent factual inaccuracies, such as anatomically incorrect structures or unrealistic clinical environments, which risk misleading viewers about how procedures or workplaces actually look (Cevik et al., 2025; Noel, 2024). Research on human detection of AI generated images suggests that these distortions can be difficult to spot, with people often operating near chance when asked to distinguish real from synthetic visuals and showing systematic cognitive tendencies such as impostor bias and automation bias in their judgments (Casu et al., 2025). In this review, we define content bias as distortions in what is shown and how accurately educational content, contexts, and relationships are portrayed. We examine how the included studies evaluate the realism, correctness, and narrative framing of AI generated images, and how they link these content level issues to learners' understanding of disciplines, institutions, and social issues in education.

1.3. Research framework and questions

Despite growing concern about algorithmic bias in generative AI, empirical work on AI generated text-to-image in education remains fragmented across disciplines, tools, research methods, and contexts. Studies differ in how they define and operationalize bias, which social categories they attend to, and how explicitly they connect visual representation to educational aims such as inclusion, identity, and critical AI literacy (Alon & Levkovich, i-an press). Existing studies also vary in how rigorously they document prompts, model settings, sampling strategies, and analytic procedures. Many focus on a single tool or small sets of prompts, offer limited transparency about how images were selected or coded, or give little space to students' and teachers' perspectives on bias. As a result, it is difficult to compare findings across studies, assess the robustness of their conclusions, or identify consistent methodological standards for examining bias in AI generated images for education. The purpose of this review is therefore to map and critically analyze how bias and representation are addressed in the emerging body of literature on educational uses of AI generated text-to-image tools.

We adopt a multidimensional framework of representational bias in AI generated educational imagery. In this review, representational bias is used as an umbrella construct to capture systematic distortions in AI generated images that are pedagogically consequential. Within this framework, we distinguish between demographic representation, cultural and religious representation, and content related distortions that concern epistemic accuracy, contextual specificity, or genericization of artifacts, spaces, and environments (Currie et al., 2024; Vartiainen et al., 2025). To operationalize this framework, we synthesize evidence across six dimensions: gender; racial, ethnic, and socio economic identity; cultural and religious affiliation; age; body and (dis)ability; and content bias (Elsharif et al., 2025).

These six dimensions were selected because they captured the most recurrent forms of representational bias identified across the included studies and because each reflects a pedagogically meaningful aspect of how educational images can shape visibility, belonging, and normative expectations. The dimensions are treated jointly because educational images can function as exemplars that normalize particular identities, roles, and settings while marginalizing others, with implications for inclusion and critical AI literacy (Reed et al., 2023; Vartiainen et al., 2025). This framework guides the review design by structuring the research questions, the coding scheme, and the interpretation of findings.

To make this framework explicit and operational, we summarize the six representational dimensions, their working definitions, and how they were coded in the included studies in Table 1.

Based on the multidimensional framework of representational bias (Table 1), we ask the following research questions:

RQ1. What methodological approaches and educational contexts characterize studies that examine bias and representation in AI generated text-to-image in education?
RQ2. How is AI generated text to image currently used in educational contexts?
RQ3. What forms of representational bias are documented in AI generated text-to-image used in educational settings across the framework dimensions (gender; race, ethnicity, and socio economic identity; culture and religion; age; body and disability; and content)?
RQ4. How do educators and students perceive, interpret, and respond to biases in AI generated text-to-image?
RQ5. How are the framework dimensions of bias patterned by publication year, study design, and AI image generator?

Table 1
Multidimensional framework of representational bias in AI generated educational imagery.

Framework dimension	Working definition	What we coded	Why it matters in educational visuals
Gender	Skewed gender visibility and gendered roles.	Presence, roles, stereotypes.	Normalizes who is seen as expert or authority (Currie et al., 2024; Vargas-Veleda et al., 2025).
Race, ethnicity, and socio economic identity	Skewed racialized cues, skin tone, and socio economic contexts.	Skin tone, roles, setting resources.	Shapes belonging cues and default norms (Elsharif et al., 2025; Currie et al., 2024).
Culture and religion	Defaulting, flattening, or misrepresenting cultural or religious cues.	Symbols, language, place cues.	Privileges dominant codes and erases specificity (Avis et al., 2025; Elsharif et al., 2025).
Age	Skewed visibility and stereotyping by age.	Age presence, roles, stereotypes.	Signals who is visible and valued (Gisselbaek et al., 2024).
Body and (dis) ability	Erasure or stereotyping of body and disability diversity.	Disability cues, assistive tech, body diversity.	Normalizes narrow standards of bodies (Currie et al., 2024; Vargas-Veleda et al., 2025).
Content bias	Misleading or decontextualized depictions affecting meaning.	Inaccuracy, genericization, mismatch.	Increases epistemic risk through plausible visuals (Cevik et al., 2025; Noel, 2024).

2. Methods

2.1. Study design

This systematic review was conducted in accordance with the methodological framework of the Joanna Briggs Institute (JBI) for systematic reviews (Peters et al., 2015) and followed the PRISMA-ScR reporting guidelines (Page et al., 2021). The objective was to map empirical research examining *AI-generated images* within educational contexts, with particular attention to bias, representation, inclusivity, and their instructional use.

2.2. Eligibility criteria

The eligibility criteria were defined using the PICOS framework (Population, Intervention/Exposure, Comparator/Context, Outcomes, Study Characteristics). These criteria guided both title/abstract screening and full-text assessment. Table 2 summarizes the inclusion and exclusion criteria.

2.3. Search strategy

A comprehensive search strategy was iteratively developed in consultation with all participating authors. Initial exploratory searches were conducted to identify relevant keywords and controlled vocabulary related to AI-generated images, education, and bias. The final systematic search was conducted in the following databases: Scopus, Web of Science, ERIC, EBSCO (education collections), Taylor & Francis Online, Wiley Online Library, and the ACM Digital Library (for peer-reviewed conference papers). Google Scholar was used only for citation chasing and to identify additional peer-reviewed articles.

Search strings combined terms for AI-generated images and text-to-image models with terms for educational settings and bias/representation. The syntax and subject headings were adapted for each database. A generic version of the search string was: (“text-to-image” OR “AI-generated image*” OR “image generation”) AND (education OR school* OR universit*) AND (bias OR stereotype* OR representation).

We included only peer-reviewed journal articles and full conference papers from the ACM Digital Library, published in English from 2021 onward, to capture research conducted during the rapid development of generative AI models. Although the search covered studies published from 2021 onward, all included studies were published between 2023 and 2025, likely reflecting the public emergence and early educational uptake of widely accessible generative AI tools during this period.

2.4. Screening and study selection

All citations identified in the search were imported into Covidence for automated de-duplication and screening management. Screening was conducted in three stages: (1) Pilot screening: 10% of titles/abstracts were independently screened by the authors to refine inclusion criteria and calibrate decision-making; (2) Title and abstract screening: The first author screened all records, and two additional authors screened a random 10 % subset for reliability. Inter-rater agreement ($\kappa = 0.39$) fell within the fair-to-moderate range, consistent with expectations for a systematic review; (3) Full-text screening: The first author assessed the full texts; 20% were evaluated independently by two additional authors. Discrepancies were resolved through discussion until a consensus was reached. The reasons for exclusion were systematically documented in Covidence.

2.5. Data extraction

A structured data extraction form was developed and piloted prior to use. Two authors independently extracted data from a subset of studies to ensure a shared interpretation of concepts and consistency in coding.

Table 2
Inclusion and exclusion criteria.

	Inclusion	Exclusion
Study characteristics	• Any empirical study design (quantitative, qualitative, mixed-methods, design-based) examining AI-generated images in an educational context. • Peer-reviewed journal articles or peer-reviewed conference papers. • Full text in English. • Published from 2021 onward.	• Review articles, editorials, conference abstracts, and preprints not yet accepted for publication. • Non-empirical papers (editorials, commentaries, purely theoretical). • Purely technical/model papers without educational context. • Grey literature (theses, reports, book chapters) that are not peer-reviewed.
Population	• Learners (any age) engaged in educational activities using AI-generated images. • Educators, instructional designers, or school/university staff using or designing AI-generated images for teaching and learning. • Educational materials or platforms explicitly targeting learners or educators.	• General public or consumer samples with no educational aim. • Studies in which participants are students but the task is not educationally framed. • Marketing, advertising, or entertainment uses of AI images without a learning context.
Comparator/ Context	• Use of AI-generated or AI-manipulated images as part of teaching, learning, assessment, or educational content. • The study addresses representation, bias/stereotypes, diversity, inclusivity, or fairness in these images.	• Studies using only non-AI images (e.g., video, VR, immersive technologies, non Gen-AI tools). AI-generated images used only decoratively (e.g., logos, backgrounds) with no analysis of content. • Technical image studies focused solely on realism, resolution, or accuracy, with no link to representation/bias and no educational use. • Comparisons only between AI models, datasets, or algorithms (e.g., bias metrics) without educational tasks, populations, or materials.

Following the consensus, the first author extracted data from all included studies, with secondary checks performed on 20 % of the sample.

The extracted variables included the study characteristics: (1) publication year, country, population, educational context, study design, sample size, and platform used; (2) intervention/exposure characteristics: type of AI-generated images, purpose within the educational setting, learning task, and theoretical framing (if applicable); and (3) representation and bias measures: how studies conceptualized or operationalized bias, stereotypes, inclusivity, and fairness. All extracted data were exported from Covidence to Excel for organization and preparation for the narrative synthesis, and a detailed overview of each included study is provided in Appendix A.

A total of 3005 studies were retrieved from the database search. After removing 363 duplicates, 2642 records proceeded to title and abstract screening. During this stage, 2611 studies were excluded for reasons including lack of relevance (n = 2199), not meeting the population criteria (n = 41), unsuitable study design (n = 82), not meeting the required study characteristics (n = 94), and studies that did not involve AI-generated or AI-manipulated images in a bias-related context (n = 195). Following the full screening process, 31 studies met all eligibility criteria and were included in the final review. Fig. 1 illustrates the PRISMA flow chart.

2.6. Data analysis and synthesis

We developed a framework-based coding scheme aligned with the five research questions and six representational dimensions that were selected based on recurring patterns across the included studies and their pedagogical relevance in educational settings. For each included study, we coded: (1) study characteristics (year, educational level, participant group, study design, and AI platform); (2) educational uses of AI-generated images (e.g., objects of analysis, resources in learning and teaching activities, part of expert work); (3) documented forms of bias and representation along six dimensions: gender, racial and ethnic identity, cultural and religious affiliation, age, body and (dis)ability, and socio-economic status, with an additional “other/unspecified” category; and (4) educators' and students’ perceptions and responses to these biases.

The first author applied the coding scheme to all studies using Excel. The other authors independently coded a subset of studies to check for shared interpretation of the categories. Discrepancies were resolved through discussion and minor refinements to the coding definitions. Descriptive statistics were used to summarize the study characteristics and distribution of representation dimensions. A narrative synthesis was then conducted, organizing the findings according to the five research questions and highlighting recurrent patterns and methodological gaps. Table 3 summarizes this study's coding scheme.

3. Findings

Findings are organized according to the coding scheme. We first describe the characteristics of the reviewed studies and then present the results for each research question.

3.1. Characteristics of the reviewed studies

The final corpus comprised 31 studies published between 2023 and 2025, despite the broader search window beginning in 2021. The number of publications increased sharply over time: three in 2023, eight in 2024, and 20 in 2025, indicating that educational research on text-to-image systems is very recent and rapidly expanding.

In terms of educational level, 12 studies were situated explicitly in K–12 settings (e.g., physics, social studies, Bible teaching, and AI literacy for children), while the remaining 19 studies focused on higher education, teacher education, or general educational imagery (e.g., medicine, nursing, architecture, design, and university literature courses).

Regarding participant groups, 12 studies worked only with AI-generated images and had no human participants. Among the 19 studies that involved people, three included K–12 students, four focused on teachers, nine involved university students, two on university faculty, and one on other K-12 experts (e.g., specialists co-designing an AI-literacy tool for children).

Study designs were grouped into four methodological categories: image-focused studies without human participants (n = 12); qualitative and design-based studies with participants (n = 11); mixed-methods studies (n = 5); and experimental or other quantitative designs (n = 3).

Across the 31 studies, researchers used a wide range of generative image tools. DALL·E (including DALL·E 2 and DALL·E 3, often accessed via ChatGPT) was the most common (n = 16), followed by Midjourney (n = 11), Stable Diffusion (n = 7), and Adobe Firefly (n = 6). Smaller numbers of studies used Bing Image Creator/Copilot (n = 4), Leonardo (n = 3), and other generators such as Playground, Krea, or Lexica, typically in comparative or benchmarking designs. Table 4 summarizes the characteristics of the reviewed studies.

3.2. Educational uses of AI generated images

Across the 31 studies, AI-generated images were used in three main ways: as objects of analysis (n = 12), as resources in learning and

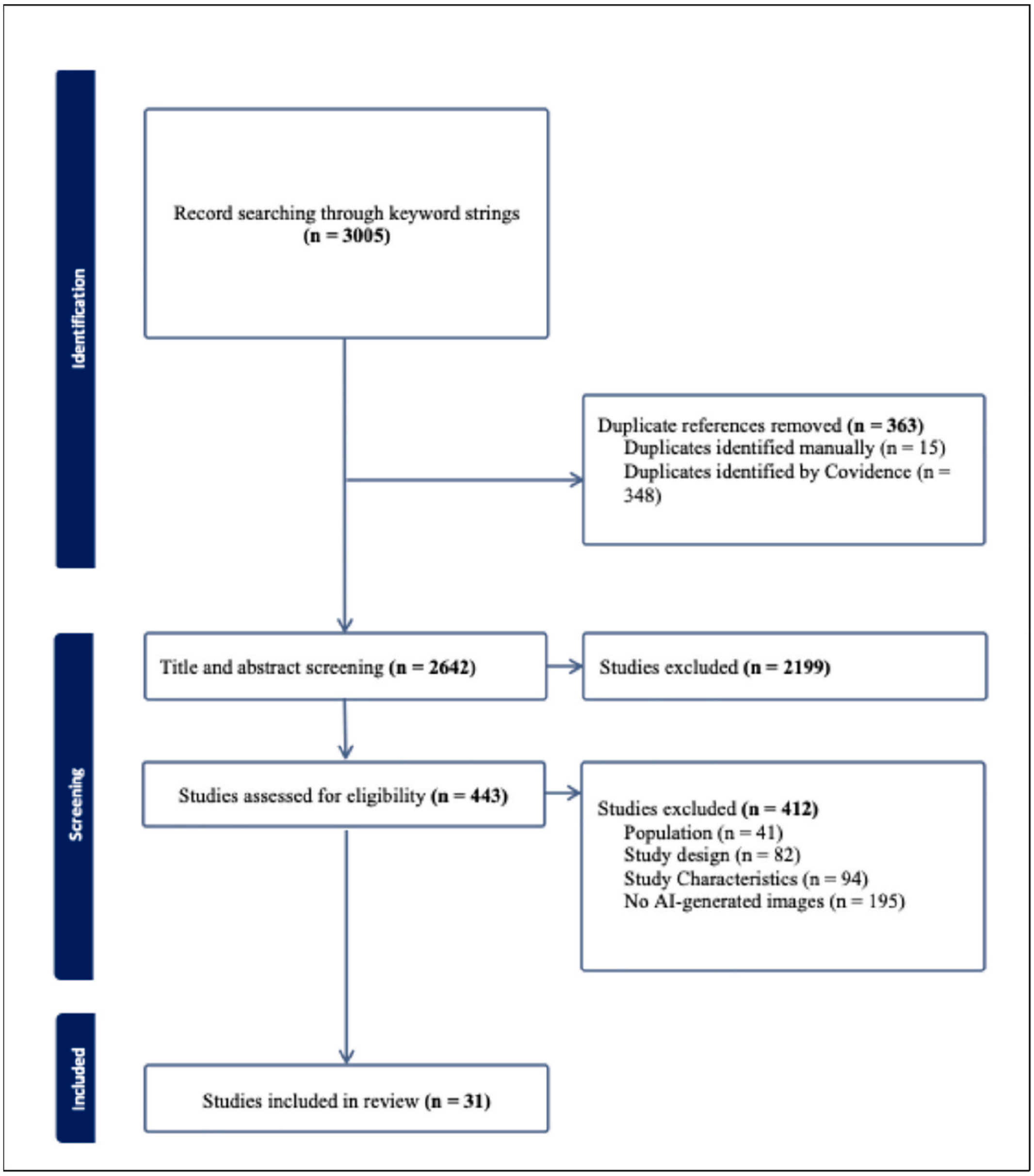

Fig. 1. PRISMA flow chart for selecting articles for this review study.

Table 3
Coding scheme.

Category	Codes/sub-codes	Description
Study characteristics	Year, educational level, participant group, study design, AI tool.	Basic descriptive info for mapping the field.
Educational use of images	Objects of analysis, resources in learning and teaching activities, part of expert work.	How AI images are embedded in teaching/learning.
Bias dimensions	Gender; race/ethnicity/SES; culture/religion; age; body/(dis) ability; content; other.	Which representational patterns are reported.
Response to bias	Teacher/students perspectives.	How actors perceive, interpret, or act on the biases.

teaching activities (n = 16), and as part of expert work (n = 3). In the objects-of-analysis studies (n = 12), researchers generated images using text-to-image tools and then examined these images as potential educational materials. Examples included depictions of medical students and patients, university academics, scientists, classrooms, and biblical or historical scenes. The images were not used with students; instead, the focus was on how text-to-image systems might shape future teaching resources and identify biases or inaccuracies before classroom adoption. Eight of these studies were situated in higher education, and four were in K–12 contexts.

In learning and teaching activities (n = 16), AI-generated images were embedded directly into educational tasks. Ten of these studies were conducted in higher education and six in K–12 settings. University students used text-to-image tools to design architectural and urban spaces, create heritage- or literature-related visuals, explore scientific and medical concepts, or represent professional identities. K–12 students used AI-generated images to illustrate physics phenomena, construct characters and story scenes, or examine how different occupations and social groups are portrayed. In some studies, images were pre-generated and presented in videos, slides, or worksheets; in others, students authored prompts themselves, viewed the outputs, and reflected on how the images influenced their understanding.

Finally, expert work (n = 3) involved teachers, faculty, or other

Table 4
Characteristics of the included studies (n = 31).

Characteristic	Category	n	% of studies
Educational level	K–12	12	39%
	Higher education/teacher education	19	61%
Participant group	No human participants	12	39%
	K–12 students	3	10%
	University students	10	32%
	Teachers	4	13%
	University faculty	3	10%
	Other educational experts	1	3%
Study design	Qualitative image/content analysis	12	39%
	Qualitative/design-based	11	36%
	Mixed-methods	5	16%
	Experimental/quantitative	3	10%
AI tools[a]	DALL·E (any version, incl. via ChatGPT)	16	–
	Midjourney	11	–
	Stable Diffusion (incl. online variants)	7	–
	Adobe Firefly/Firefly	6	–
	Bing Image Creator/Copilot	4	–
	Leonardo	3	–
	Other (Playground, Krea, Lexica, Craiyon)	1 each	–

[a] Multiple tools could be used in a single study; therefore, percentages are not reported for tool use.

education specialists who engaged with AI-generated images in professional development or design contexts. Two of these studies were located in K–12 education and one in higher education. In this group, text-to-image systems were used to prototype culturally responsive visuals, co-design AI literacy tools for children, and test how reliably local cultures, professions, and social issues were represented. Here, images primarily supported reflection on curriculum, representation, and ethics, rather than serving as direct student-facing content.

3.3. Bias identified in the studies

Across the 31 included studies, bias in AI text-to-image was the rule rather than the exception. Different studies highlighted overlapping dimensions of bias, which we group here as gender; race/ethnicity/socio-economic status (SES); culture/religion; age; body/ability; content/epistemic issues; and other-process or system level biases. The following findings show that generative AI imagery in education is shaped by multiple, intersecting biases: who appears and in what roles; which cultures, places, and socio-economic contexts are made visible; how bodies and abilities are depicted; and how accurately disciplinary knowledge is represented.

3.3.1. Gender

Gender bias was identified in 12 of the 31 included studies. These studies showed systematic over-representation of men in high-status or expert roles and concentration of women in caring, support, or feminised roles (Campbell et al., 2025; Collyer-Hoar et al., 2025; Cooper & Tang, 2024; Currie et al., 2024, 2025a, 2025b; Ringvold et al., 2024; Thomson et al., 2025; Tianshu and Sukumara, 2025; Vartiainen et al., 2025; Vartiainen & Tedre, 2023; Yetisensoy, 2025). For example, “university academic” and “medical student” prompts produced large majorities of male figures, while early childhood education prompts generated mainly feminine women in aprons as educators and formally dressed men as leaders, visually linking care with femininity and leadership with masculinity (Currie et al., 2024, 2025a; Thomson et al., 2025).

3.3.2. Race, ethnicity, and SES

Race/ethnicity and skin-tone bias appeared in 16 of the 31 studies. These papers documented a consistent dominance of light-skinned or white figures, with medium or dark skin tones markedly under-represented (Bakkum et al., 2024; Campbell et al., 2025; Currie et al., 2024, 2025a, 2025b; Fareed et al., 2024; Heruti, 2025; Nascimiento et al., 2024; Reed et al., 2023; Ringvold et al., 2024; Thomson et al., 2025; Tianshu and Sukumara, 2025; Vartiainen et al., 2025; Vartiainen & Tedre, 2023; Yetisensoy, 2025; de Souza et al., 2024). In several professional domains (e.g., academics, medical imaging professionals, early childhood educators), dark-skinned characters were rare or absent, and no visible ethnic diversity appeared in key roles (Currie et al., 2024, 2025a; Thomson et al., 2025).

Socio-economic bias in educational settings was described in a subset of these 16 studies. Images of school science and classroom environments depicted spacious, highly resourced labs, advanced equipment, and orderly rows of desks, aligning science learning with elite or high-SES schools and marginalizing under-resourced contexts (Cooper & Tang, 2024; Vartiainen et al., 2025).

3.3.3. Culture and religion

Cultural and religious bias was reported in 10 of the 31 studies (Bakkum et al., 2024; Chrostowski & Najda, 2025; Condorelli & Berti, 2025; Dakamsih & Alkayid, 2025; Fareed et al., 2024; He et al., 2025; Heruti, 2025; Huh et al., 2025; Vartiainen & Tedre, 2023; Yetisensoy, 2025). These studies showed that generative models defaulted to Western, Anglocentric, or Christian visual codes even when prompts invoked non-Western contexts. Examples included European-looking religious figures in Bible teaching, English text in signage instead of local languages, Western humanitarian symbols (e.g., Red Cross instead of Red Crescent), and exaggerated cultural motifs that turned everyday settings into caricatures. Heritage and local-architecture work showed that famous global sites were rendered convincingly, whereas small towns, local costumes, and vernacular buildings were replaced by generic or “Disneyfied” scenes (Condorelli & Berti, 2025; Fareed et al., 2024; Yetisensoy, 2025).

3.3.4. Age

Age bias was explicitly identified in four of the 31 studies (Bakkum et al., 2024; Currie, Hewis, & Wheat, 2025; Vartiainen et al., 2025; Vartiainen & Tedre, 2023). These studies reported a default towards young adult or middle-aged figures. Older adults appeared infrequently or in stereotyped roles (for example, the “old scientist” trope), and case vignettes generated by large language models tended to produce young, healthy patients unless age was specified with great precision (Bakkum et al., 2024; Vartiainen et al., 2025). As a result, late-life stages and intergenerational configurations were weakly represented in educational imagery.

3.3.5. Body and (dis)ability

Body size and (dis)ability bias was present in 3 of the 31 studies (Bakkum et al., 2024; Currie, Hewis, & Wheat, 2025; Vartiainen et al., 2025). Where coded, images overwhelmingly depicted normatively thin, non-disabled bodies; visible disabilities and assistive devices were absent. In clinical vignettes, generated narratives defaulted to healthy-weight patients and linked higher BMI with particular occupations or lifestyles in stereotyped ways (Bakkum et al., 2024). Studies on academics and STEM figures similarly described an erasure of larger-bodied and disabled people from AI-generated scenes (Currie, Hewis, & Wheat, 2025; Vartiainen et al., 2025).

3.3.6. Content and epistemic bias

Content and epistemic bias, in which images looked plausible but were scientifically, historically, or contextually inaccurate, was documented in 10 of the 31 studies (Bakkum et al., 2024; Basty et al., 2025; Campbell et al., 2025; Condorelli & Berti, 2025; Fareed et al., 2024; He et al., 2025; Nascimiento et al., 2024; Noel, 2024; Rakhshi et al., 2025; de Souza et al., 2024). In anatomy, medicine, chemistry, and physics, models produced attractive visuals with missing, distorted, or impossible structures, or with misleading conceptual metaphors (Campbell et al., 2025; Nascimiento et al., 2024; Noel, 2024; Rakhshi et al., 2025). Heritage and architecture work identified hallucinated or generic buildings and costumes that were substituted for real local features (Condorelli & Berti, 2025; Fareed et al., 2024). Clinical vignette generation revealed inconsistent BMI, medications, and allergies within the same case, which compounded representational biases with technical inaccuracies (Bakkum et al., 2024).

3.3.7. Other biases: stylistic, algorithmic, and process-level

Stylistic, algorithmic, and process level biases were discussed in 15 of the 31 studies (Bakkum et al., 2024; Basty et al., 2025; Chrostowski & Najda, 2025; Collyer-Hoar et al., 2025; Cooper & Tang, 2024; Doumpioti & Huang, 2024; Huh et al., 2025; Iranmanesh & Lotfabadi, 2025; Kang et al., 2025; Nascimento et al., 2024; Reed et al., 2023; Vartiainen et al., 2025; Vartiainen & Tedre, 2023; Yetisensoy, 2025; de Souza et al., 2024). Architectural and design education studies described a drift towards spectacular landmark aesthetics and homogenized visual styles driven by training data, which ignored local regulations, material constraints, or craft traditions (Iranmanesh & Lotfabadi, 2025; Kang et al., 2025; Vartiainen & Tedre, 2023). Studio based and design thinking courses reported that human evaluators and juries interacted with AI generated outputs in ways that amplified existing preferences for particular styles or visual cues (Basty et al., 2025; Doumpioti & Huang, 2024; Reed et al., 2023).

Several of these studies used generative AI as a probe for bias. They designed AI literacy or critical media activities in which students analyzed stereotypes, experimented with prompts, and surfaced patterns that reflected the underlying training data (Collyer-Hoar et al., 2025; Reed et al., 2023; Vartiainen et al., 2025).

3.4. Educators' and students' responses to bias in AI generated images

Across the 31 studies, 19 studies that involved human participants described how educators or students responded when they encountered bias or limitations in AI-generated images. Overall, participants moved between cautious use, active critique, and uncritical acceptance, with these modes often co-existing within the same study.

In 10 studies, educators and students treated generative AI text-to-image tools as useful but limited instruments. They valued these tools for speeding up ideation, visualization, and preparation of teaching materials, and for supporting students who struggle with drawing or design. Yet they did not trust the images as final products when cultural representation or technical accuracy mattered (Basty et al., 2025; Doumpioti & Huang, 2024; Fareed et al., 2024; Huh et al., 2025; Kang et al., 2025; Vartiainen & Tedre, 2023). Participants emphasized the need for human oversight and positioned AI as a complement rather than a replacement for existing visual resources, sometimes editing, curating, or replacing AI outputs when images felt stereotypical, generic, or culturally inappropriate.

Eight studies reported explicit critical engagement with bias. In these studies, students, teachers, or educational experts actively named stereotypes and discussed who was shown and who was missing in the images, linking these patterns to training data or platform design (Collyer-Hoar et al., 2025; Heruti, 2025; Reed et al., 2023; Ringvold et al., 2024; Vartiainen et al., 2025). Participants critiqued gendered and racialized portrayals, questioned idealized depictions of professional roles, and reflected on how biased visuals might shape children's or students' aspirations. In studies that included structured teaching activities, children's ability to explain that AI reflects its training data and can misrepresent groups increased over time, although misconceptions about how the system works did not disappear completely (Heruti, 2025; Vartiainen et al., 2025).

Seven studies highlighted over-trust in AI outputs or limited spontaneous recognition of bias. Students and children sometimes accepted AI images as realistic or authoritative, focused mainly on conceptual correctness or aesthetic appeal, and overlooked inaccuracies, generic depictions of place, or culturally flattened global scenes (Condorelli & Berti, 2025; Dakamsih & Alkayid, 2025; de Souza et al., 2024; Romanik & Wegner, 2025). In these studies, biased or homogenized visual defaults risked becoming normalized unless educators intervened. Students rarely raised representational issues without explicit prompting, which underscores the need for structured pedagogical scaffolding when AI images are incorporated into teaching and assessment (Condorelli & Berti, 2025; Fareed et al., 2024). Two studies described emotional responses to biased or mismatched imagery, where students found AI-generated images and videos engaging yet also reported discomfort or confusion when visuals clashed with their lived experiences or expectations of working conditions (Dakamsih & Alkayid, 2025; Reed et al., 2023). These emotional reactions sometimes opened space for deeper reflection but did not on their own guarantee that participants would recognize or challenge bias without guided discussion.

Overall, these 19 studies show that AI-generated images do not automatically foster critical awareness. Without explicit framing and guidance, students tended either to treat images as convenient but opaque tools or to over-trust them as realistic depictions. Where educators deliberately used biased or partial outputs as teaching moments, students were more likely to articulate how AI can misrepresent people, places, and professions.

3.5. Patterns of bias by year, study design, and tool

To address RQ5, we examined how the seven bias types (six main dimensions and an 'other' category) are distributed across publication year, study design, and AI tool. This allows us to see whether particular forms of bias cluster in specific periods, methodological approaches, or AI tools. Because individual studies often reported multiple bias dimensions and used several AI tools, totals across tables may exceed 31. Each cell represents the number of studies mentioning a given bias type within that category rather than unique studies: *Bias* × *Year* counts each bias theme by publication year; *Bias* × *Design* counts each bias theme by methodological type; *Bias* × *Tool* counts each bias theme by image generator used.

3.5.1. Bias by publication year

Across publication years, the total number of reported bias instances increased from 11 in 2023 to 14 in 2024 and 41 in 2025, indicating a rapid expansion of attention to bias in AI generated images. Race, ethnicity, and socioeconomic status were the most consistently examined category, rising from 3 mentions in 2023 to 4 in 2024 and 9 in 2025. Gender, culture and religion, content, and other biases were discussed only sporadically in 2023 and 2024, but each appeared between 6 and 8 times in 2025, showing a broadening of the bias agenda beyond race. Age and body or ability remained comparatively under examined across all years, with 2 and 4 mentions respectively. Overall, these patterns reflect both growth in the volume of studies and a clearer articulation of multiple intersecting bias dimensions over time (see Fig. 2).

3.5.2. Bias by study design

Across research methods, most bias instances were identified in image analysis studies (29 of 69) and qualitative studies with human participants (23 of 69), with fewer in mixed methods (13) and experimental or quantitative designs (4). Race, ethnicity, and SES bias was documented 14 times, mainly via image analysis (8) and qualitative work (3). Gender bias documented 12 times, again concentrated in image analysis (7) and qualitative work (4). Culture and religion bias (11 instances) appeared primarily in qualitative (5) and mixed methods studies (4), while content bias (10) and "other" biases (15) were distributed across all approaches but remained highly skewed toward image analysis and qualitative designs. Age (3) and body or ability (4) biases were rarely examined and did not appear in experimental work.

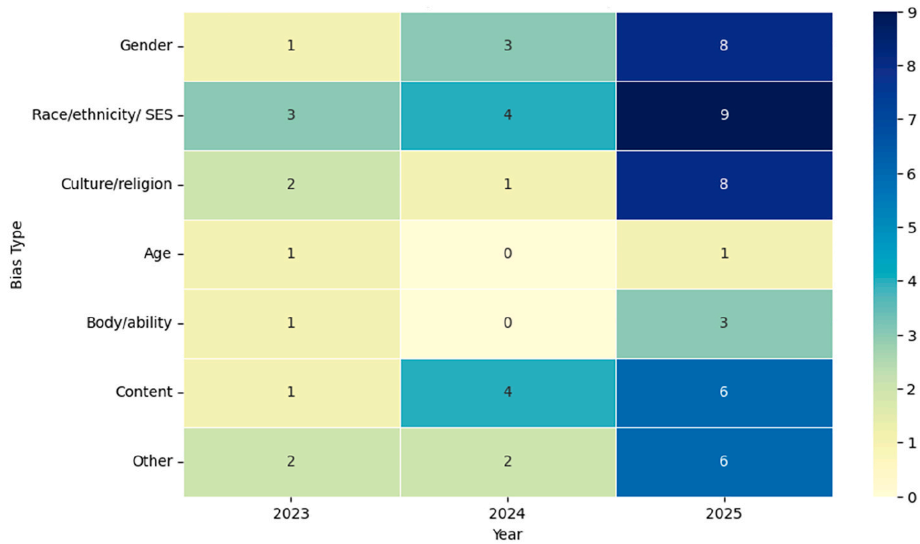

Fig. 2. Bias by publication year.

Overall, current evidence on bias in educational AI images is built almost entirely on descriptive and interpretive designs rather than experimental or strongly quantitative approaches (see Fig. 3).

3.5.3. Bias by AI-tool

Across AI tools, bias instances clustered most around DALL·E (29 of 95) and Midjourney (20), with smaller but non-trivial counts for other tools. Content bias was the most frequently observed category overall (22 instances), spread across all tools and especially visible in DALL·E (4), Stable Diffusion (4), Firefly (3), and Bing/Copilot (3). Race, ethnicity, and SES bias (19 instances) and gender bias (17 instances) were also widespread, with DALL·E again carrying the largest share (7 race, 6 gender). Culture and religion bias appeared 13 times, most often in DALL·E (5) and Midjourney (4), while body and (dis)ability (4) and age (2) were rarely noted and only in a subset of tools. Overall, the mapping suggests that core representational and content biases are not confined to a single generator but recur across multiple platforms, with higher counts for DALL·E and Midjourney reflecting both their popularity and their central role in educational studies (see Fig. 4).

Overall, the interaction patterns described in this section suggest that research on bias in AI images in education is both expanding and maturing, but still uneven. Over time, studies have moved beyond a narrow focus on a single bias to a more layered view that includes race and ethnicity, gender, culture, and content, while dimensions such as age and body or (dis)ability remain rarely examined. Methodologically, most evidence comes from image audits and qualitative work with learners and educators, so we know about how bias appears and how people interpret it, but much less about causal mechanisms or the impact of specific interventions. Comparing tools shows that similar forms of bias surface across several popular generators, rather than being confined to one platform, which points to shared training data and design assumptions. The interaction analysis highlights both a growing critical agenda and clear blind spots, revealing the need for more diverse methods and a broader set of bias dimensions in future studies.

4. Discussion

This review synthesizes empirical evidence from 31 studies published between 2023 and 2025, capturing a formative period in the uptake of text-to-image tools in education. Because text-to-image

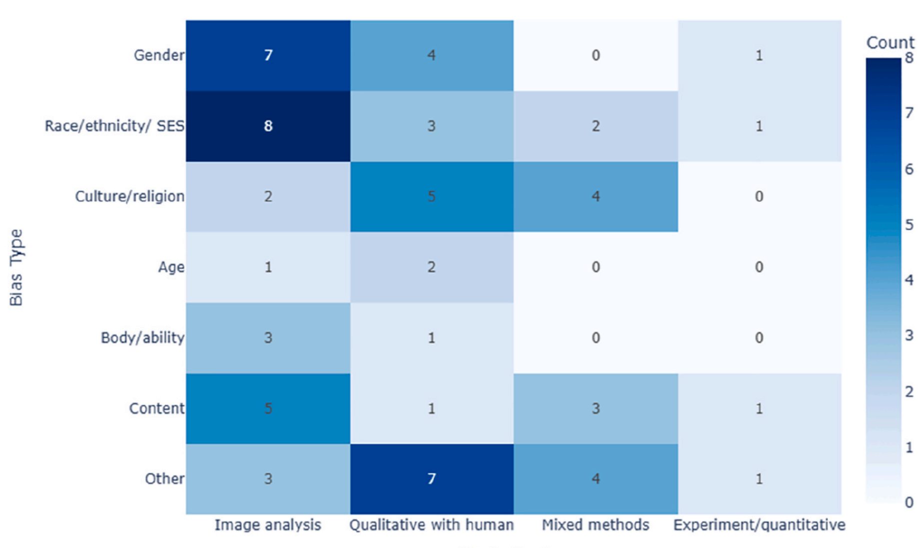

Fig. 3. Bias by study design*. *Category counts can exceed the number of included studies because studies may be coded into multiple categories and categories are not mutually exclusive.

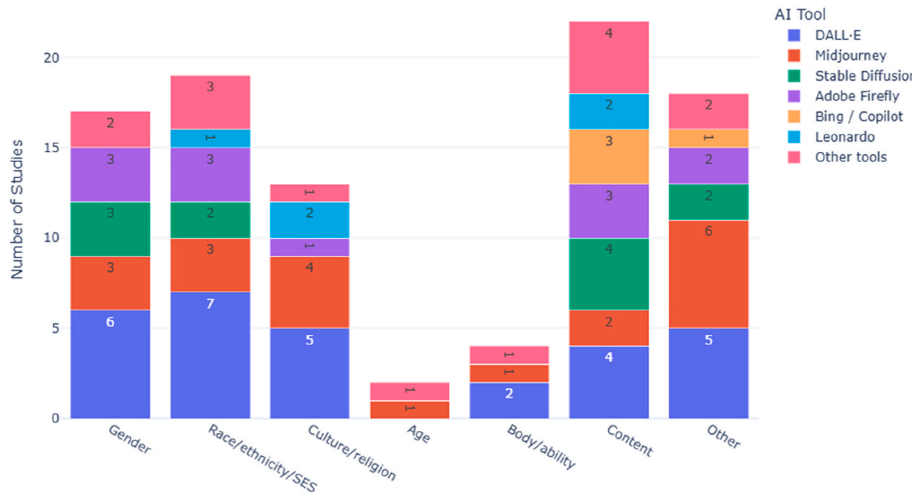

Fig. 4. Bias by AI tool. *"Other tools" includes smaller or mixed-use generators (e.g., Playground, Krea, Lexica, Craiyon, Imagine AI) and the ChatGPT-only vignette studies; Category counts can exceed the number of included studies because studies may be coded into multiple categories and categories are not mutually exclusive.

models changed rapidly during this period, patterns documented in earlier studies may reflect different architectures and mitigation strategies than later systems, which affects how generalizable specific observations are across time. Across K-12 and higher education, the literature documents rapid classroom experimentation with tools such as DALL·E and Midjourney, alongside growing attention to bias and representation across six framework dimensions. The evidence also points to a persistent gap between the ease of producing visually plausible AI generated images and the pedagogical infrastructures needed to evaluate them. AI generated images can reproduce and, in some cases, intensify representational distortions, including omission, stereotyping, and narrowed defaults. As text-to-image tools move from occasional experimentation to routine classroom use, educators face a dual challenge: leveraging generative capacity for teaching and learning while systematically identifying and addressing representational risks.

Most included studies document representational patterns and user perceptions, so the implications discussed here concern pedagogical risk and classroom use conditions rather than measured learning outcomes. The discussion is organized around three implications that emerge consistently across the included studies. First, recurring representational defaults and omissions shape who is made visible in AI generated educational images across the six framework dimensions. Second, the visual plausibility of AI generated images creates epistemic risks when outputs are used as instructional exemplars despite inaccuracies or decontextualized depictions. Third, the reviewed evidence indicates that critical visual AI literacy in education requires structured evaluation routines that make representational goals explicit and reviewable in everyday use.

A related consideration concerns disciplinary focus. In studies where prompts primarily generate images of people, representational problems are most often documented as demographic skew, omission, and stereotyping across identity dimensions. In contrast, in design, architecture, and place based contexts where outputs often foreground environments, artifacts, or scenes rather than human figures, representational concerns more commonly appear as stylistic homogenization, generic imagery, or loss of local cultural and historical specificity. In the framework used in this review, these patterns are captured primarily under the culture and religion and content dimensions, reflecting that representational distortion in educational images can concern both who is depicted and how settings and contexts are rendered.

4.1. Representational defaults shape who is visible in education

While gender and racial or ethnic skews are the most frequently documented issues in AI generated images used for education, the synthesis also points to a more structurally important pattern: omission. Across the reviewed studies, age and body and (dis)ability are among the least examined dimensions, and when they are examined they are consistently under represented in AI generated outputs. Older adults appear infrequently, often in narrow tropes, and late life roles are weakly represented unless age is specified with unusual precision (Vartiainen et al., 2025; Vartiainen & Tedre, 2023). Visible disability markers and assistive devices are similarly rare or absent across depictions of learners, educators, and professionals, reinforcing an implicit norm of able bodied participation (Bakkum et al., 2024; Currie, Hewis, & Wheat, 2025).

This pattern aligns with Tuchman's (1978) concept of symbolic annihilation, where systematic non representation functions as marginalization. In educational contexts, omission has direct relevance because AI generated images are increasingly used as exemplars for roles and settings. When students and educators generate visuals for prompts such as scientist, teacher, classroom, or academic, repeated exposure to image sets dominated by young, non disabled bodies can implicitly define who is expected to belong in these roles and who remains outside the visual field. This concern is reinforced by the broader representational defaults documented across the corpus, including the dominance of light skinned figures and narrow occupational portrayals in educational and professional prompts (Currie et al., 2024, 2025a; Ringvold et al., 2024; Thomson et al., 2025).

A practical implication follows for educational use and for research design. Omission should be treated as an explicit indicator of representational bias when evaluating AI generated images for teaching and learning, alongside stereotyping and skew. This requires that studies and classroom activities document who is visible across a set of generated images, which identity cues were specified or left implicit in prompts, and whether the resulting visuals align with the stated inclusion goals of the task.

4.2. Visual plausibility and epistemic risk in AI generated images

A second implication concerns the recurrent tension between the visual plausibility of AI generated images and their factual, disciplinary, or contextual accuracy. Across the reviewed studies, models produced outputs that appeared coherent and authoritative while containing important distortions. In line with the review framework, we treat these issues as epistemic and contextual distortions in AI generated images rather than as demographic bias, although they can interact with demographic representation in classroom use. This pattern was documented in STEM and medical learning activities, where visually convincing diagrams or depictions included missing or incorrect structures, and in history and architecture contexts where scenes and buildings were rendered in generic forms that reduced local specificity (Campbell et al., 2025; Condorelli & Berti, 2025; Fareed et al., 2024; Nascimiento et al., 2024; Rakhshi et al., 2025). These findings place content bias alongside demographic representation as a core educational concern, because learners often engage images as explanatory objects and as anchors for disciplinary understanding.

The reviewed evidence also suggests that plausibility can shape how learners and educators evaluate outputs. Several studies report that users frequently judge images by surface coherence and aesthetic appeal, and that over-trust can occur when realistic appearance is taken as a proxy for correctness (Condorelli & Berti, 2025; de Souza et al., 2024; Romanik & Wegner, 2025). Conceptually, this pattern aligns with well documented tendencies to treat algorithmic outputs as epistemic authorities and to rely on them as defaults when they appear confident or realistic, even when verification is warranted. In our corpus, this is consistent with evidence on automation bias and related judgment effects in evaluating AI generated images, which helps clarify why plausibility can override accuracy checks in classroom use (Casu et al., 2025).

In classroom settings, this combination creates a specific epistemic risk: inaccurate or decontextualized images can be adopted as teaching materials or learner artifacts with limited scrutiny, particularly when time constraints and novelty effects reduce careful verification. The implication for educational practice is that text-to-image activities require explicit verification routines that treat AI generated images as provisional representations. Effective routines include comparing outputs to authoritative sources, identifying where an image functions as evidence versus illustration, and documenting prompt details and tool settings so that limitations can be discussed and traced in reflective analysis (de Souza et al., 2024; Ringvold et al., 2024).

4.3. *Critical visual AI literacy as structured classroom practice*

A third implication concerns how text-to-image activities are framed pedagogically. Across the reviewed studies, learner and educator engagement with bias and representation varied with the degree of instructional scaffolding. When guidance was limited, students often treated AI generated images as realistic depictions and did not consistently question representational defaults or omissions (Dakamsih & Alkayid, 2025; Romanik & Wegner, 2025). When activities incorporated guided critique, comparison tasks, and reflective discussion, learners were more likely to identify recurring representational patterns and to articulate limitations in system outputs (Collyer-Hoar et al., 2025; Heruti, 2025; Ringvold et al., 2024). Overall, the synthesis suggests that prompt writing competence alone does not ensure critical interpretation of AI generated images.

An additional contextual consideration concerns developmental stage. The reviewed studies include both K-12 and higher education settings, and some findings suggest that younger learners may be more likely to treat AI generated images as realistic depictions and to rely on them with limited verification (de Souza et al., 2024; Vartiainen et al., 2025). In higher education contexts, several studies describe more instrumental use and, when supported by critique prompts or comparison tasks, more explicit reflection on limitations and bias (Basty et al., 2025; Doumpioti & Huang, 2024; Fareed et al., 2024; Huh et al., 2025; Kang et al., 2025). However, the current evidence base does not yet support firm conclusions about consistent developmental differences, because few studies directly compare age groups using comparable tasks, prompts, and evaluation criteria. This remains an important direction for future work, particularly for designing age appropriate scaffolds for critical visual AI literacy.

Over, these findings support an applied view of critical visual AI literacy as a set of routinized practices that make representational goals explicit and reviewable in everyday use. This involves designing tasks that require students to evaluate outputs across the framework dimensions, attend to omission alongside stereotypes, and justify image selection decisions rather than relying on the first plausible output. It also involves transparency norms, including recording prompts, noting tool or model versions when available, and documenting selection criteria, so that classroom critique can examine how representational outcomes emerge through the interaction between tool behavior and user choices (Ringvold et al., 2024; Reed et al., 2023; Vartiainen et al., 2025). For teacher education, the implication is that professional development should extend beyond tool familiarization to include structured evaluation routines for AI generated images that align with equity goals and with disciplinary expectations for accuracy and contextual integrity (Alon & Levkovich, *i-an press*).

The review also offers a theoretical contribution by advancing a multidimensional framework of representational bias in AI generated educational images that is operationalized for both analysis and educational use. By consolidating six recurring dimensions that are often treated separately, the framework provides a coherent lens for evaluating classroom images and for designing instructional interventions. In teacher education and professional development, this lens can be translated into a rubric that supports consistent review of AI generated images across dimensions, making inclusion goals explicit and auditable. Such a rubric can guide educators to examine presence and omission, role distributions, cultural and contextual cues, and content related distortions, while documenting prompts and selection criteria. In this way, the framework provides shared language that supports more comparable research and more targeted professional development around text-to-image use in educational settings.

Taken together, the reviewed evidence indicates that bias and representation in AI generated images for educational settings operate through recurring defaults and omissions, through epistemic risks tied to plausibility, and through variation in whether classroom use is supported by structured critical practices. These implications point to the need for clearer pedagogical guidance and for research designs that connect representational patterns and user interpretations to instructional decisions, classroom norms, and learning tasks across diverse educational contexts.

4.4. *Limitations and future studies*

This systematic review has several limitations that should be considered when interpreting the findings. First, because visual representation and bias are culturally and linguistically situated, the focus on English language publications likely constrained the geographic and epistemic scope of the evidence base. This may have limited coverage of text-to-image use in non-Western educational contexts, as well as locally grounded analyses of cultural and religious representation, language cues, and place specific meanings in AI generated images. Future reviews should incorporate multilingual search strategies and regional databases to better capture diverse educational settings and representational standards.

Second, the rapid evolution of generative AI models creates challenges for stability and reproducibility. Tool providers frequently update model architectures, training data, and safety filters, and many platforms do not provide transparent versioning or access to historical model states. As a result, specific outputs documented in the reviewed studies may not be reproducible, and patterns observed at one point in time may shift as models and policies change. Future studies should therefore report prompts, settings, sampling procedures, and tool versions when available, and should adopt designs that can assess how representational outcomes change across updates, including longitudinal audits and repeated prompt protocols.

Third, the distribution of attention across framework dimensions remains uneven. The reviewed literature provides clearer evidence on gender and race or ethnicity, while age, disability, religion, cultural specificity, and socio economic cues are less frequently examined and are often operationalized inconsistently. This limits the ability to compare results across studies and to assess intersectional patterns that are central to educational equity. Future research should extend auditing beyond single dimensions by using designs that code multiple dimensions simultaneously, clarify decision rules for ambiguous cues, and examine how prompt wording, task framing, and image selection practices shape representational outcomes in educational settings. Finally, more work is needed to connect representational patterns to educational processes, including how students and educators interpret AI generated images, how these images are selected and used in instruction, and which scaffolds and evaluation routines support more critical and inclusive use across K-12 and higher education contexts.

5. Conclusion

The rapid integration of text-to-image tools in education offers substantial creative possibilities, and this review shows that it also introduces persistent representational risks. Across the included studies, AI generated images commonly display recurring defaults, omissions, and stereotyped portrayals across the framework dimensions, alongside content related distortions that can appear visually plausible. These patterns matter for education because classroom images function as exemplars of who belongs in disciplinary roles and because visually convincing depictions can be adopted as instructional materials even when they are inaccurate or decontextualized. These findings position text-to-image use as a pedagogical and governance issue, calling for structured evaluation routines and critical visual AI literacy in everyday practice. The multidimensional framework advanced in this review offers a coherent basis for teacher education, professional development, and the development of rubrics to assess AI generated images by making representational goals explicit and reviewable.

CRediT authorship contribution statement

Lilach Alon: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Dorit Hadar Shoval:** Writing – review & editing, Writing – original draft. **Inbar Levkovich:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Review of included studies

Coding note: The categories reported in the Appendix, follow the review framework and were applied to enable consistent synthesis across studies. We coded each study based on how the authors framed and operationalized the phenomenon under review, and we did not reclassify a study's claims beyond what was supported by its reported methods and results. For studies focused on cultural or historical accuracy, generic representation, or non-human imagery, we coded these findings under content bias or culture and religion only when the study explicitly examined systematic distortions in representation, context, or specificity in AI generated images relevant to educational use. When a study discussed these issues as accuracy or fidelity rather than bias, this was retained in the extraction notes, and the coding reflects grouping for synthesis rather than a change to the study's original framing.

	Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
1.	Bakkum et al. (2024)	Using artificial intelligence to create diverse and inclusive medical case vignettes for education	Higher education	Medical education	No human. 8-prompt workflow	Qualitative image/content analysis	ChatGPT-3.5	objects of analysis	Body, age, ethnic/cultural, and intersectional bias: ChatGPT defaulted to young, healthy-weight, often white patients and produced stereotypes (e.g., “French” wine connoisseur, “cook” with high BMI). Attempts to randomize for diversity yielded implausible profiles that misrepresented real-world intersectionality, and technical errors (BMI, medications, allergies) added further content bias.	NA
2.	Basty et al. (2025)	Exploring higher education faculty insights on generative AI in creative courses	Higher education	Art education	University faculty (N = 53)	Mixed method	DALL·E, Midjourney, Stable Diffusion, Adobe Firefly, Bing, Copilot	Expert work	Representation bias and lack of diversity: faculty worried that skewed training data lead to stereotypical, inaccurate, or exclusionary images, which can misrepresent people and contexts and undermine inclusion and equity in creative courses.	Faculty were enthusiastic but cautious: they valued text-to-image tools for ideation, visualization, and preparing materials, yet expressed strong concern that biased or misleading images could misguide students and erode

(continued on next page)

(*continued*)

	Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
										originality and craft, especially if students over-rely on AI.
3.	Boddepalli & Solanke, 2025	Evaluating the role of generative AI and ChatGPT in pharmacology education: A medical student perspective	Higher education	Medical education	University students (N = 140)	Experimental/ quantitative study	ChatGPT-3.5	Resources in learning/ teaching activities	The authors, drawing on prior evidence of gender and ethnic bias in medical and educational imaging, urge caution in using generative AI tools.	Students voiced concerns about accuracy and potential misuse of AI-generated images (e.g., that they might be misleading or used inappropriately), but their responses to representation-related bias were not specifically examine.
4.	Campbell et al. (2025)	Generative artificial intelligence images for instruction: A content analysis	Higher education	General higher education	No human. 291 AI-generated images.	Qualitative image/content analysis	DALL-E2 Stable Diffusion. Adobe Firefly	Objects of analysis	Gender, racial/skin-tone, and stereotypical bias, plus accuracy issues: 21.3% of images showed gender bias (53.4% men vs 45.2% women), 18.9% race-related bias, 19.6% skin-color bias, and 17% stereotypical "scientist" tropes (lab coats, glasses). Anatomical errors were common (e.g., distorted hands in 22.3% and eye disfigurement in 17.9% of images), raising concerns about visual validity.	NA
5.	Chrostowski and Najda (2025)	From verse to vision: Exploring AI-generated religious imagery in Bible teaching	K-12	Bible studies	No human. 4 AI-generated images	Qualitative image/content analysis	DALL·E 3 in ChatGPT-4o	Objects of analysis	Religious and cultural bias: AI-generated Bible images simplify or omit key religious symbols and meanings and reproduce Western Christian stereotypes, such as a European-looking Jesus with little visible Jewish or historical context.	NA
6.	Collyer-Hoar et al. (2025)	Experts unite, kids delight: Co-designing an inclusive AI literacy educational tool for children	K-12	AI literacy	Education experts (N = 7)	Qualitative study with educators or learners	Stable Diffusion XL	Expert work	Gender bias: generative image models can reproduce gender stereotypes and shape children's ideas about "who can be what." The co-design team anticipates that certain traits and hobbies will push the AI toward more "male" or "female" appearing characters, revealing how models encode stereotypically feminine and masculine attribute.	Education experts involved in co-design explicitly acknowledge the risk that biased visuals may influence children's views of roles and identities. They treat this as a design challenge to be surfaced and managed, not ignored.
7.	Condorelli and Berti (2025)	Creativity and awareness in co-creation of art using artificial intelligence-based systems in heritage education	Higher education	Heritage education	University students (N = 16)	Mixed method	Firefly, DALL·E 3	Resources in learning/ teaching activities	Content and cultural bias: AI, trained on global datasets, accurately depicted famous sites but produced generic or "fake" images of small villages, missing local architecture and costumes and overusing clichéd heritage elements (food, customs). The authors also note ethical/ legal bias (copyright and restricted data) and sustainability bias (high energy use at odds with	Students often trusted AI outputs too much, sometimes failing to notice or correct inaccurate or generic images of local heritage. Their tendency to accept AI-generated visuals at face value exposed how dataset bias can quietly shape understanding when users do not critically question the images.

(*continued on next page*)

(*continued*)

	Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
									heritage education's sustainability goals).	
8.	Cooper and Tang (2024)	Pixels and pedagogy: Examining science education imagery by generative artificial intelligence	K-12	Science education	No human. 8 AI-generated images	Qualitative image/content analysis	DALL·E 3 integrated in ChatGPT-4	Objects of analysis	Strong stereotypical and socio-economic bias: AI-generated scenes are dominated by white lab coats, goggles, microscopes, dense formulas, and well-equipped labs that evoke elite, high-SES schools, marginalizing under-resourced contexts. Although there is some gender and ancestry diversity, it appears within narrow visual conventions, with nearly all educators still in white lab coats (often with glasses).	NA
9.	Currie et al. (2024)	Gender bias in generative artificial intelligence text-to-image depiction of medical students	Higher education	Medical education	No human; 47 DALL·E images.	Qualitative image/content analysis	DALL·E 3 integrated in ChatGPT; Adobe Firefly	Objects of analysis	Strong gender and racial/ethnic bias: for individual "medical student" prompts, 92% of figures were male and all were light-skinned; group images were ~57% male and over 90% light-skinned. DALL·E 3 thus heavily over-represents white men compared to the majority-female Australian medical student population, while "nursing student" prompts produced roughly 50/50 male–female images, indicating the bias is specific to medical student.	NA
10.	Currie, Hewis, and Wheat (2025)	Gender and ethnicity representation of university academics by generative artificial intelligence using DALL-E 3	Higher education	Medical education	No human. 81 AI-generated images	Qualitative image/content analysis	DALL·E 3 integrated in ChatGPT-4	Objects of analysis	Strong gender, racial/ethnic, body, and ability bias in depictions of university academics: 82% male vs 16% female, about 94% light-skinned with almost no dark-skinned figures, and no visible disabilities or larger body types. Bias is strongest in individual portraits, which are almost all white men.	NA
11.	Currie, Hewis, Hawk, and Rohren (2025)	Gender and Ethnicity Bias of Text-to-Image Generative Artificial Intelligence in Medical Imaging, Part 2: Analysis of DALL-E 3	Higher education	Medical education	No human. 120 AI-generated images	Qualitative image/content analysis	DALL·E 3 integrated in ChatGPT-4	Objects of analysis	Strong gender and racial/ethnic bias: about 57% of medical imaging professionals were male and 91% light-skinned, with dark-skinned characters rare and no visible disabilities. Nurses were more often female and physicians more often male, but overall depictions remained heavily skewed toward light-skinned men.	NA
12.	Dakamsih and Alkayid (2025)	When stories get a visual voice: AI-generated images, a visual pedagogy for	Higher education	Literature education	University students (N = 30)	Mixed methods	Midjourney	Resources in learning/	Cultural bias and interpretive flattening: AI-generated visuals often imposed Western-	Students found AI images engaging and emotionally involving, with women more

(continued on next page)

(*continued*)

	Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
		enhancing student engagement and critical analysis in world literature						teaching activities	looking landscapes and imagery on world literature texts, limiting cultural nuance and narrowing possible readings of the stories. By presenting a single, polished visual, the images tended to fix one dominant interpretation, reducing openness and multiplicity of meaning	drawn to symbolic elements and men to literal or philosophical ones. Yet 75% treated the visuals as definitive, and over half felt they sometimes constrained meaning or clashed with the cultural context.
13.	de Souza et al. (2024)	Visualising relativity: Assessing high school students' understanding of complex physics concepts through AI-generated images	K-12	Physics education	High school students (N = 12)	Qualitative study with educators or learners	Bing	Resources in learning/ teaching activities	Content bias: the conceptual quality of AI-generated images depends strongly on students' prompts. Vague or tangential prompts produce misleading or conceptually empty images, while even students with good understanding may obtain distorted visuals if they cannot clearly translate their ideas into text.	Students evaluated the images mainly in terms of conceptual correctness and clarity, adjusting prompts when outputs did not match their intended ideas. They did not explicitly discuss representational bias, but their struggles and revisions show an emerging awareness that AI images can misrepresent what they mean if the prompt is unclear.
14.	Doumpioti and Huang (2024)	Collaborative design with generative AI and collage: Enhancing creativity and participation in education	Higher education	Design education	University students (N = 13)	Qualitative study with educators or learners	Midjourney	Resources in learning/ teaching activities	Human evaluative bias in design assessment suggests that AI-supported workflows might complement, rather than replace, human judgement in evaluation.	Students saw text-to-image AI as helpful but limited: it sped up ideation, supported collaboration, and let non-experts join in, but outputs often missed their mental image and felt lower in quality or human touch. They treated GenAI as a complement rather than a replacement for sketching, models, and CAD.
15.	Fareed et al. (2024)	Exploring the potentials of artificial intelligence image generators for educating the history of architecture	Higher education	Architecture education	University students (N = 30)	Mixed methods	Leonardo AI	Resources in learning/ teaching activities	Cultural and content bias: AI-generated images of historical architecture were often inaccurate, overly polished, "Disneyfied", and sometimes historically wrong. Experts warned that popular outputs frequently reflect Western/orientalist views of non-Western heritage, risking distorted understandings of local history and architecture.	Students were enthusiastic and often treated AI reconstructions as realistic, while educators were more cautious, consistently pointing out inaccuracies, authenticity problems, and the need to validate AI images before using them in teaching.
16.	He et al. (2025)	"Carlitos the curious caterpillar": Exploring teacher-AI co-creation of culturally responsive educational materials for young learners	K-12	Early childhood education	Teachers (N = 5)	Qualitative study with educators or learners	DALL·E 3 integrated in ChatGPT-4	Expert work	Cultural bias and misrepresentation risk: teachers repeatedly encountered AI outputs that were not culturally accurate or locally grounded, which confirmed that text-to-image systems cannot be trusted to represent culture responsibly without human intervention.	Teachers did not trust AI to represent culture on its own. They actively edited AI outputs for scientific correctness, age appropriateness, and cultural relevance, and deliberately inserted their own cultural details (local names, places, Spanish words) instead of accepting AI-generated versions of these elements.

(*continued on next page*)

(*continued*)

	Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
17.	Heruti (2025)	Shifting between image and text with an AI image generator: From personal experience to multicultural awareness and recognition bias	Higher education	Multicultural education	University students (N = 12)	Mixed methods	Playground, Leonardo	Resources in learning/teaching activities	Strong cultural and ethnic bias: AI images defaulted to Western, white, idealized figures, with multicultural diversity underrepresented. The systems amplified stereotypical cultural symbols (e.g., Middle East = desert, Japan = kimono, Israel = religious iconography), reinforcing reductive and essentialized portrayals of cultural groups.	Students explicitly identified and critiqued these cultural and representational biases, developing heightened critical awareness of how algorithmic systems construct, distort, and simplify cultural identities, and reflecting more deeply on race, ethnicity, and representation.
18.	Huh et al. (2025)	Students' perceptions of generative AI image tools in design education: Insights from architectural education	Higher education	Architecture education	University students (N = 42)	Qualitative study with educators or learners	Midjourney	Resources in learning/teaching activities	Bias is addressed indirectly through concerns that AI-generated images lack contextual, cultural, and human-centered awareness. Students noted that the tools can produce culturally insensitive or inappropriate designs and propagate structural or spatial errors, reflecting limited understanding of real users and places.	Students responded with cautious distrust, accepting GenAI for early ideation but rejecting it for final-stage work because of its questionable design choices, shallow contextual understanding, and risk of biased or inappropriate outputs without human oversight.
19.	Iranmanesh and Lotfabadi (2025)	Critical questions on the emergence of text-to-image artificial intelligence in architectural design pedagogy	Higher education	Architecture education	University students (N = 41) Faculty (N = 6)	Qualitative study with educators or learners	DALL·E, Midjourney	Resources in learning/teaching activities	Algorithmic and contextual bias: because training sets over-represent iconic, dramatic architectural images, broad prompts produce stereotyped, landmark-like, symmetrical, fantasy buildings that ignore local context, regulations, gravity, and human scale.	Students and faculty recognized that these set images encourage generic, contextless design thinking. They viewed the bias toward spectacular, collage-like forms as a serious limitation for architectural pedagogy, reinforcing the need for critical distance and human judgement.
20.	Kang et al. (2025)	AI-assisted design communication in ceramic-crafts education: Investigating image generation tools for concept representation and pedagogical practice	Higher education	Craft education	University students (N = 12) Faculty (N = 9)	Experimental/quantitative study	Stable Diffusion v1.5 ControlNet	Resources in learning/teaching activities	Stylistic and technical bias: AI tools tend to generate homogenized styles and occasionally physically impossible or misleading ceramic forms, which can distort students' sense of feasibility and originality.	Students appreciated AI tools for enriching design but could be overly trusting of technically impossible forms. Faculty valued AI-assisted prototypes yet expressed concern that style homogenization.
21.	Nascimiento et al., 2024	Enhancing AI responses in chemistry: Integrating text generation, image creation, and image interpretation through different levels of prompts	Higher education	Chemistry education	No human. 7 generative AI systems	Qualitative image/content analysis	ChatGPT 3.5/4.0, Bard, Bing, Adobe Firefly, Leonardo.AI, DALL·E	Objects of analysis	Content (epistemic/disciplinary) bias and inaccuracy: Text-to-image tools frequently produced chemically incorrect, confusing, or oversimplified images, and vision models sometimes generated confident but wrong explanations, risking reinforcement of common chemistry misconceptions.	NA
22.	Noel (2024)	Evaluating AI-powered text-to-image generators for anatomical illustration: A comparative study	Higher education	Medical education	No human. 9 AI-generated images.	Qualitative image/content analysis	Bing, Stable Diffusion, Craiyon V3	Objects of analysis	Content bias and inaccuracy: All evaluated systems generated images that were visually appealing but anatomically incorrect,	NA

(*continued on next page*)

(*continued*)

	Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
									with missing or wrong structures in the skull, heart, and brain, making them unreliable as anatomical representations.	
23.	Rakhshi et al. (2025)	Can AI illustrate science? A comparative benchmarking study of text-to-image artificial intelligence models for scientific communication	Higher education	Science education	No human. 120 AI-generated images	Qualitative image/content analysis	DALL·E v3, Midjourney v5.2, Stable Diffusion v2.1.	Objects of analysis	Content bias: General-purpose text-to-image models prioritised aesthetic appeal over scientific correctness: only about half the images reflected the intended concept, very few were contextually related, and none of the 120 images were fully accurate scientifically.	NA
24.	Reed et al. (2023)	AI image-generation as a teaching strategy in nursing education	Higher education	Nursing education	University students (N = 4)	Qualitative study with educators or learners	Midjourney	Resources in learning/ teaching activities	Gender and racial bias: prompts with nurse produced almost exclusively white, female nurses; male nurses appeared only when explicitly requested and were still mostly white and conventionally presented. Some uniforms were historically styled, reinforcing an outdated, traditional image of nursing.	Students noticed and named these patterns as biases and stereotypes, contrasting the idealized, calm, beautifully lit nurse–patient scenes with their expectations of real working conditions and using the images to reflect on their own fears and desired professional identity.
25.	Ringvold et al. (2024)	The AI generative text-to-image creative learning process: An art and design educational perspective	K-12	Art and design education	Teachers (N = 4)	Qualitative study with educators or learners	Midjourney; also, Krea, Lexica, Stable Diffusion Online, Adobe Firefly	Resources in learning/ teaching activities	Gender and racial/ethnic default biases: when gender or ethnicity are not specified, images tend to default to white male figures; prompts like superhero mostly yield male heroes; and vague prompts frequently generate stereotypical young women. More detailed descriptors can partially counter these defaults but still rely on fixed, categorical labels.	Teachers recognized these defaults as problematic and stereotypical, noting that AI tools can both democratize visualization and simultaneously reproduce narrow ideals of who gets represented as powerful, heroic, or visible in visual culture.
26.	Romanik & Wegner, 2025	Effects of videos with AI-generated images to provide students with authentic and diverse insights into STEM occupations	K-12	STEM education	Gifted children (N = 137) ages 7–12	Experimental/ quantitative study	Midjourney, DALL·E 2	Resources in learning/ teaching activities	STEM-related stereotypes (e.g., traits and images associated with STEM professionals) rather than coding the AI images themselves in detail. Despite using AI-generated visuals intended to be authentic and diverse, stereotype scores showed mixed patterns: some traits shifted in a positive direction in the experimental group, while other stereotypes became more extreme rather than more moderate.	Gifted children who watched the AI-based videos reported learning something new and having their interest sparked, but showed no significant gains in STEM interest, self-efficacy, or aspirations versus controls; stereotype measures suggested the videos alone did not reliably reduce, and sometimes even reinforced, stereotypical views.
27.	Thomson et al. (2025)	Exploring bias in artificial intelligence: Stereotypes and gendered narratives in digital imagery of early childhood educators	K-12	Early childhood education	No human. Prompts entered into ChatGPT	Qualitative image/content analysis	DALL·E 3 integrated in ChatGPT	Objects of analysis	Gender and racial/ethnic bias: GenAI images cast early childhood educators mainly as white, feminine women in aprons doing nurturing, hands-on	NA

(*continued on next page*)

(*continued*)

Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
								work, and leaders as white men in corporate clothing with passive children, aligning care with femininity and leadership with white masculinity. The near-total lack of racial/ethnic diversity and the contrast between active children with women vs passive with men narrow who "belongs" in ECE and who "looks like" a leader.	
28. Tianshu & Sukumaran, 2025	Integrating AI-based image generation in design education: A mixed-methods quasi-experimental study on academic performance improvement	K-12	Design education	No human. 8 AI-generated images	Qualitative image/content analysis	DALL-E 3	Objects of analysis	Gender and racial/ethnic bias: educators are mostly white, feminine women in aprons doing nurturing, hands-on work, while leaders are more often male, formally dressed authority figures. Children are active with female educators but passive with male educators, and the minimal racial/ethnic diversity reinforces a narrow, Eurocentric and de-professionalized image of the ECE workforce.	NA
29. Vartiainen et al. (2025)	Enhancing children's understanding of algorithmic biases in and with text-to-image generative AI	K-12	AI literacy	Elementary and middle school children (N = 213)	Qualitative study with educators or learners	Midjourney	Resources in learning/teaching activities	Stereotypical bias: firefighters, bullies and gamers were mostly depicted as boys; scientists as old, messy-haired men; poor people as dirty old men; Jesus as white; and woman as a narrow, idealized body type. These reflect gender, racial/ethnic, body, and socio-economic bias in Midjourney's default outputs.	Initially, most children either could not explain the bias or attributed it to how the AI thinks. After the teaching sequence, more students correctly described bias as arising from imbalanced training data and proposed concrete data-level fixes, indicating improved algorithmic bias understanding, though some misconceptions persisted.
30. Vartiainen & Tedre, 2023	Using artificial intelligence in craft education: Crafting with text-to-image generative models	K-12	Craft education	Teachers (N = 15)	Qualitative study with educators or learners	Midjourney	Resources in learning/teaching activities	Algorithmic and cultural bias: training data can misrepresent or under-represent groups, repeat narrow stylistic norms, and reinforce Anglocentric and stereotyped imagery, with platforms shaping which visual styles and identities become dominant.	Teachers were enthusiastic but wary: they valued text-to-image tools for ideation and for supporting pupils who struggle to draw, yet worried that biased, decontextualized designs and platform-driven aesthetics could erode craft thinking, ignore real-world constraints, and reproduce unequal power over visual culture.
31. Yetisensoy (2025)	Validity challenge in GenAI models: Evaluating the validity of content generated by text-to-image models in the context of social studies education	K-12	Social studies education	Teachers (N = 5)	Qualitative study with educators or learners	DALL-E 3	Resources in learning/teaching activities	Linguistic and cultural bias: DALL·E 3 often used English instead of Turkish, substituted Western symbols (e.g., Red Cross for Red Crescent), and exaggerated cultural	Teachers judged some images as valid and useful but systematically flagged others as factually wrong, Westernized, or culturally distorted, recognizing that the

(continued on next page)

(*continued*)

Authors	Title	Education level	Discipline	Participants group	Study design	AI tools	Use of images	Bias dimensions	Response to bias
								elements (e.g., a giant Turkish carpet covering a modern classroom), distorting everyday Turkish contexts.	model's representations could mislead students about local symbols, environments, and everyday life.

References

Alon, L., & Levkovich, I. (*in press*). Information literacy in the age of generative tools: Development and validation of the AI information literacy scale (AILIS). Computers in Human Behavior: Artificial Humans.

Avis, G. R., Nocco, S., Zhao, A. Q., & Cremaschi, M. (2025). Guiding AI-Assisted persona generation with Hofstede's cultural dimensions: A mixed-methods evaluation. *Behaviour & Information Technology*, 1–29. https://doi.org/10.1080/0144929X.2025.2581258. .

Bakkum, M. J., Hartjes, M. G., Piët, J. D., Donker, E. M., Likic, R., Sanz, E., de Ponti, F., Verdonk, P., Richir, M. C., van Agtmael, M. A., & Tichelaar, J. (2024). Using artificial intelligence to create diverse and inclusive medical case vignettes for education. *British Journal of Clinical Pharmacology, 9*(3), 640–648.

Basty, R., Kropczynski, J., & Halse, S. (2025). Exploring higher education faculty insights on generative AI in creative courses. *Journal of Information Technology Education: Research, 24*. https://doi.org/10.28945/5546

Bender, S. M. (2023). Coexistence and creativity: Screen media education in the age of artificial intelligence content generators. *Media Practice and Education, 24*(4), 351–366. https://doi.org/10.1080/25741136.2023.2204203. .

Bian, C., Wang, X., Huang, Y., Zhou, S., & Lu, W. (2025). Effects of AI-generated images in visual art education on students' classroom engagement, self-efficacy and cognitive load. *Humanities and Social Sciences Communications, 12*, 1548. https://doi.org/10.1057/s41599-025-05860-2

Boddepalli, D., & Solanke, B. P. (2025). Evaluating the role of generative AI and ChatGPT in pharmacology education: A medical student perspective. *International Journal of Life Sciences Biotechnology and Pharma Research, 14*(5), 598–605. https://doi.org/10.69605/ijlbpr_14.5.2025.104. .

Campbell, L. O., Lambie, G. W., Hayes, B. G., & Hartshorne, R. (2025). Generative artificial intelligence images for instruction: A content analysis. *TechTrends: Linking Research & Practice to Improve Learning*, 1–13. https://doi.org/10.1007/s11528-025-01129-2

Casu, M., Guarnera, L., Zangara, I., Caponnetto, P., & Battiato, S. (2025). A (Mid)journey through reality: Assessing accuracy, impostor bias, and automation bias in human detection of AI-generated images. *Human Behavior and Emerging Technologies, 2025*(1), Article 9977058. https://doi.org/10.1155/hbe2/9977058. .

Cevik, J., Seth, I., & Rozen, W. M. (2025). Artificial intelligence-generated images in skin cancer education: An experimental study. *Australasian Journal of Plastic Surgery, 8*(1), Article 118181. https://doi.org/10.34239/ajops.118181

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*, 43. https://doi.org/10.1186/s41239-023-00411-8

Chrostowski, M., & Najda, A. J. (2025). From verse to vision: Exploring AI-generated religious imagery in Bible teaching. *Religions, 16*(8), 1051. https://doi.org/10.3390/rel16081051

Collyer-Hoar, G., Castellani, A., Guluzade, L., Tomczyk, B., Bilal, H., & Rubegni, E. (2025). Experts unite, kids delight: Co-designing an inclusive AI literacy educational tool for children. In *Proceedings of the 24th interaction design and children* (pp. 852–857). https://doi.org/10.1145/3713043.3731495

Condorelli, F., & Berti, F. (2025). Creativity and awareness in co-creation of art using artificial intelligence-based systems in heritage education. *Heritage, 8*(5), 157. https://doi.org/10.3390/heritage8050157

Cooper, G., & Tang, K.-S. (2024). Pixels and pedagogy: Examining science education imagery by generative artificial intelligence. *Journal of Science Education and Technology, 33*(4), 556–568. https://doi.org/10.1007/s10956-024-10104-0

Currie, G., Currie, J., Anderson, S., & Hewis, J. (2024). Gender bias in generative artificial intelligence text-to-image depiction of medical students. *Health Education Journal, 83*(7), 732–746. https://doi.org/10.1177/00178969241274621. .

Currie, G., Hewis, J., Hawk, E., & Rohren, E. (2025b). Gender and ethnicity bias of text-to-image generative artificial intelligence in medical imaging, part 2: Analysis of DALL-E 3. *Journal of Nuclear Medicine Technology, 53*(2), 162–168. https://doi.org/10.2967/jnmt.124.268359

Currie, G., Hewis, J., & Wheat, J. (2025a). Gender and ethnicity representation of university academics by generative artificial intelligence using DALL-E 3. *Journal of Further and Higher Education, 49*(8), 1064–1078. https://doi.org/10.1080/0309877X.2025.2528788

Dakamsih, N., & Alkayid, A. (2025). When stories get a visual voice: AI-generated images, a visual pedagogy for enhancing student engagement and critical analysis in world literature. *Educational Process: International Journal, 18*. https://doi.org/10.22521/edupij.2025.18.416. .

de Souza, M. G., Won, M., Treagust, D., & Serrano, A. (2024). Visualising relativity: Assessing high school students' understanding of complex physics concepts through AI-generated images. *Physics Education, 59*(2), Article 025018. https://doi.org/10.1088/1361-6552/ad1e71

Doumpioti, C., & Huang, J. (2024). Collaborative design with generative AI and collage: Enhancing creativity and participation in education. *Data-Driven Intelligence, 2*, 445–454.

Elsharif, W., Alzubaidi, M., & Agus, M. (2025). Cultural bias in text-to-image models: A systematic review of bias identification, evaluation and mitigation strategies. *IEEE Access, 10*, 122636–122659. https://doi.org/10.1109/ACCESS.2025.3585745. .

Fareed, M. W., Bou Nassif, A., & Nofal, E. (2024). Exploring the potentials of artificial intelligence image generators for educating the history of architecture. *Heritage, 7*(3), 1727–1753. https://doi.org/10.3390/heritage7030081

Gisselbaek, M., Minsart, L., Köselerli, E., Suppan, M., Meco, B. C., Seidel, L., & Berger-Estilita, J. (2024). Beyond the stereotypes: Artificial Intelligence image generation and diversity in anesthesiology. *Frontiers in Artificial Intelligence, 7*, Article 1462819. https://doi.org/10.3389/frai.2024.1462819. .

He, K., Liu, X., Xu, Y., Bustamante, A. S., & Warschauer, M. (2025). "Carlitos the curious caterpillar": Exploring teacher-AI co-creation of culturally responsive educational materials for young learners. In *Proceedings of the 24th interaction design and children* (pp. 236–254). https://doi.org/10.1145/3713043.3727056

Heruti, V. (2025). Shifting between image and text with an AI image generator: From personal experience to multicultural awareness and recognition bias. *Multicultural Education Review, 17*(1), 59–83. https://doi.org/10.1080/2005615X.2025.2467802

Hu, Y. (2025). Generative AI, communication, and stereotypes: Learning critical AI literacy through experience, analysis, and reflection. *Communication Teacher, 39*(1), 6–12. https://doi.org/10.1080/17404622.2024.2397065

Huh, M. B., Miri, M., & Tracy, T. (2025). Students' perceptions of generative AI image tools in design education: Insights from architectural education. *Education Sciences, 15*(9), 1160. https://doi.org/10.3390/educsci15091160

Iranmanesh, A., & Lotfabadi, P. (2025). Critical questions on the emergence of text-to-image artificial intelligence in architectural design pedagogy. *AI & Society, 4*(5), 3557–3571. https://doi.org/10.1007/s00146-024-02111-x

Kang, W., Lee, H.-S., & Hong, J.-H. (2025). AI-assisted design communication in ceramic-crafts education: Investigating image generation tools for concept representation and pedagogical practice. *Computers and Education, 239*, Article 105419. https://doi.org/10.1016/j.compedu.2025.105419

Liao, C. W., Chen, H. W., Chen, B. S., Wang, I. C., Ho, W. S., & Huang, W. L. (2025). Exploring the application of text-to-image generation technology in art education at vocational senior high schools in Taiwan. *Information, 16*, 341. https://doi.org/10.3390/info16050341

Nascimento, W. J. D., Morais, C., & Girotto Junior, G. (2024). Enhancing AI responses in chemistry: Integrating text generation, image creation, and image interpretation through different levels of prompts. *Journal of Chemical Education, 101*(9), 3767–3779. https://doi.org/10.1021/acs.jchemed.4c00230

Noel, G. P. J. C. (2024). Evaluating AI-powered text-to-image generators for anatomical illustration: A comparative study. *Anatomical Sciences Education, 17*, 979–983. https://doi.org/10.1002/ase.2336

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*. https://doi.org/10.1136/bmj.n71. .

Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation, 13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Peterson, M., Delgado, C., Tang, K. S., Bordas, C., & Norville, K. (2021). A taxonomy of cognitive image functions for science curriculum materials: Identifying and creating 'performative' visual displays. *International Journal of Science Education, 43*(2), 314–343. https://doi.org/10.1080/09500693.2020.1868609. .

Rakhshi, R., Khosravi, T., Rahimzadeh, A., Shahraki, F., & Oladnabi, M. (2025). Can AI illustrate science? A comparative benchmarking study of text-to-image artificial intelligence models for scientific communication. *Journal of Clinical and Basic Research, 9*(2), 28–33. https://doi.org/10.29252/JCBR.9.2.28. .

Reed, J., Alterio, B., Coblenz, H., O'Lear, T., & Metz, T. (2023). AI image-generation as a teaching strategy in nursing education. *Journal of Interactive Learning Research, 34*(2), 369–399.

Ringvold, T. A., Strand, I., Haakonsen, P., & Strand, K. S. (2024). The AI generative text-to-image creative learning process: An art and design educational perspective. *Design and Technology Education, 29*(2), 359–379. https://doi.org/10.24377/DTEIJ.article2433

Romanik, M., & Wegner, C. (2025). Effects of videos with AI-generated images to provide students with authentic and diverse insights into STEM occupations. *American Journal of STEM Education, 5*, 15–34. https://doi.org/10.32674/edckb827

Thomson, E., Chapman, R., Cooper, G., & Cooper, Z. (2025). Exploring bias in artificial intelligence: Stereotypes and gendered narratives in digital imagery of early

childhood educators. *Gender and Education*, 1–20. https://doi.org/10.1080/09540253.2025.2546048
Tianshu, W., & Sukumaran, S. (2025). Integrating AI-based image generation in design education: A mixed-methods quasi-experimental study on academic performance improvement. *Perinatal Journal, 33*(1), 210–218. https://doi.org/10.57239/prn.25.03310028. .
Tuchman, G. (1978). Introduction: The symbolic annihilation of women by the mass media. In G. Tuchman, A. K. Daniels, & J. Benét (Eds.), *Hearth and home: Images of women in the mass media* (pp. 3–38). Oxford University Press.
Vargas-Veleda, Y., Rodríguez-González, M. del M., & Marauri-Castillo, I. (2025). Visual representations in AI: A study on the most discriminatory algorithmic biases in image generation. *Journalism and Media, 6*(3), 110. https://doi.org/10.3390/journalmedia6030110
Vartiainen, H., Kahila, J., Tedre, M., López-Pernas, S., & Pope, N. (2025). Enhancing children's understanding of algorithmic biases in and with text-to-image generative AI. *New Media & Society, 27*(9), 5342–5368. https://doi.org/10.1177/14614448241252820
Vartiainen, H., & Tedre, M. (2023). Using artificial intelligence in craft education: Crafting with text-to-image generative models. *Digital Creativity, 34*(1), 1–21. https://doi.org/10.1080/14626268.2023.2174557
Vázquez, A. F. D. C., & Garrido-Merchán, E. C. (2024). A taxonomy of the biases of the images created by generative artificial intelligence. *arXiv preprint*. arXiv:2407.01556.
Yetisensoy, O. (2025). Validity challenge in GenAI models: Evaluating the validity of content generated by text-to-image models in the context of social studies education. *Journal of Pedagogical Research, 9*(4), 81–101. https://doi.org/10.33902/JPR.202535435