

How Algorithmic Models Affect Public Attitudes and Ethical Considerations Across Different Fields

Tessa Chow¹

Received February 17, 2026

Accepted June 29, 2026

Electronic access July 31, 2026

The widespread adoption of advanced Artificial Intelligence models has integrated algorithmic decision-making into numerous fields, providing both benefits and drawbacks that intensify concerns in domains that depend on human nuance. As algorithmic systems adopt roles once reserved for humans, understanding how people evaluate and respond to decisions made by these systems becomes increasingly crucial. This paper considers how human judgment versus algorithmic models differently shape public attitudes and ethical evaluations across four domains: art, physical healthcare, mental health, and education. The overarching argument is that the acceptance of algorithmic decision-making is shaped less by technical performance and may depend on the field's degree of subjectivity and moral stakes: objective, data-driven tasks elicit greater professional acceptance, while subjective or high-stakes contexts and the importance of perceived human qualities amplify distrust elicited by lack of transparency in AI algorithms. The analysis synthesizes possible causes of algorithm aversion and draws on a literature review of research published between 2011 and 2025 through compiling quantitative surveys, experimental moral-dilemma studies, case examples, and comparative empirical findings to provide scientific context and identify recurring patterns: transparency increases trust, anthropomorphic connotations moderate acceptance, and fear of replacement and moral weight encourage negative dispositions. The review concludes with practical recommendations for hybrid designs, improved AI literacy for both the general public and field professions, and context-sensitive employment of algorithms that align use with ethical standards. Additionally, algorithms' roles continue to expand into areas essential to human welfare, and institutions should consider how people morally interpret their roles, not just the practical abilities of such models.

Index Terms — AI systems, blackbox model, anthropomorphism, algorithm aversion

AI systems: AI systems are defined as computer systems that are capable of executing tasks that typically require human approaches or intelligence (International Organization for Standardization, n.d.)¹.

Blackbox model: Much of the general public's view on the use of AI in more subjective fields stems from the lack of transparency in the models. The common term for this is a "black box model". Black box models are machine learning or statistical models in which only the outputs and inputs are observable, not the inner, complex process. The black box problem refers to the question of the trust that the public should place in deep learning models and their opaque decision-making.

Robot anthropomorphism: Robot anthropomorphism is defined as the "tendency to attribute human-like characteristics to nonhuman agents" such as robots (Roselli et al. 2025)². This personification has been seen to have a positive correlation with human acceptance and comfort

when interacting with these agents.

Algorithm Aversion (Anthropocentric Bias): As mentioned previously, algorithm aversion refers to the tendency to favor human-made decisions more than those made by algorithms or apathetic agents. Algorithm aversion is usually affected by anthropocentric bias, which is where humans and their well-being are prioritized over other stakeholders and their interests.

Introduction

The use of algorithmic systems has become much more prevalent today, with the adoption of new Artificial Intelligence (AI) models such as ChatGPT. Throughout this paper, the terms 'AI systems' (or 'AI') and 'algorithmic systems' are utilized to represent the broad spectrum of computer-driven, automated technologies assisting decision making. These algorithmic systems have become integral in some domains, yet have also raised concerns within those areas and amongst more subjective fields. Because AI can be perceived as independent moral

¹ Kent Place School, tessa.chow@yahoo.com

agents with the capability of being held accountable, moral norms may differ. In an experimental study, regardless of the ethical framework, participants rated the behavior of AI agents as more “immoral and deserving of more blame” than that of human agents (Zhang et al. 2023)³. So, understanding the ethical and social reasons behind these trends is essential for determining the benefits and drawbacks of AI’s integration into scientific fields that are crucial to people’s lives. Previous review has examined topics such as algorithm aversion and public opinions in the broad field of artificial intelligence, as well as the ethical implications in individual sectors such as healthcare and education. These existing reviews often focus on one dimension of leading factors. For example, concepts such as algorithm aversion and overall trends of anthropomorphism in AI are explored in studies such those by Zhang et al. and Maeda & Quan-Haase. Zhang et al. explores the fundamental concept of algorithm aversion, and Maeda & Quan-Haase’s literature review provides context regarding anthropomorphism without catering to a specific field. However, fewer studies evaluate studies across fields and synthesize these themes comparatively while also examining how values jointly contribute to attitudes towards algorithmic systems. This paper contributes to existing literature by providing a cross-disciplinary synthesis of AI perception across art, physical healthcare, mental healthcare, and education. Rather than examine certain variables within a single domain, this review investigates recurring patterns across fields while considering the different characteristics of the fields. By differentiating patterns across various disciplines simultaneously, this review seeks to identify broader trends that may not be visible in single-field analyses. This paper aims to provide deeper insight into the following research question: In what ways do human judgment and algorithmic models differ in their influence across fields such as physical healthcare, art, mental health, and education, and how public attitudes and ethical considerations are affected? The working hypothesis states that attitude towards these models may differ based on the moral application in the field and emotional reasoning in their use. In more subjective fields, moral weight may have a more profound impact on public view (Claim 1). Transparency, trustworthiness, and anthropomorphism are hypothesized to have a proportional relationship to people’s acceptance of AI and algorithmic systems in fields (Claim 2), while an increase in algorithm aversion and fear of replacement result in a more negative perspective on AI’s interference (Claim 3). This paper aims to synthesize research from varying fields and views on the integration and use of algorithmic models compared to human work, and investigate the motives and factors behind the discrepancies, if any, for each field. Ultimately, this paper will draw on all of these cross-field patterns and seek to inform the general sentiments about the integration of AI into human domains.

Structure of the Paper

In the following literary analysis, a consistent four-part framework is applied, with each dimension corresponding to the broader theoretical model proposed in this paper that allows for a systematic comparison of these factors. Each field discussed will include: 1) Discrepancies in public and professional opinion, the benefits of the use of AI and algorithmic systems in this area of practice, and the drawbacks of the use, 2) The transparency and trust in this use, 3) Robot anthropomorphism’s role in influencing public opinion, and 4) AI’s position as a threat versus augmentation. These 4 subsections are consistent for all topics, which discuss the fields of physical healthcare, mental health, and education, respectively.

Methodology

This paper presents a comprehensive narrative literature review that examines the differences in how human-driven ideas and algorithmic models are perceived across various factors in 4 fields: art, healthcare, mental health, and education. A thorough search was conducted using Google Scholar with keywords such as “algorithmic transparency”, “public trust”, and “anthropomorphism”. This review focused on the literature published in English from 2011 to 2025, with a geographical range of the U.S. and Europe (along with occasional contextual evidence from studies based in Asia and Nigeria). Studies were included if they 1) examined human perceptions of AI or algorithmic decision making, 2) addressed the societal or professional view on human and AI input in the relevant fields, or 3) provided psychological or ethical reasoning for this discrepancy. A total of 30 sources were selected to be included in the paper for analysis and context. Sources were included if they were peer-reviewed sources, focused on human interaction with AI or algorithm systems, included empirical data or theoretical analysis related to trust, perception, and/or ethics, and related to one or more of the selected domains (art, healthcare, mental health, education). Exclusion criteria included sources that lacked either empirical or theoretical grounding or fell outside the defined time range (post 2010). After duplicate and irrelevant sources were removed according to the PRISMA guidelines, the findings of these sources were compiled into 4 areas of interest: the public’s view as opposed to professionals’ view, the potential benefits and drawbacks from the integration of AI in the field, the role of anthropomorphism in AI on people’s judgments, and the threat or augmentation that algorithmic systems may pose in each domain. The reviewed literature examined fundamentally different types of AI intervention, including prediction, generation, and optimization. The Appendix includes Table 1A, which summarizes the primary intervention type, user group, decision task, autonomy level, outcome, and evidence context for key em-

pirical studies included in the review.

Results

Art

Discrepancies, Benefits, and Drawbacks

In a two-wave longitudinal study exploring attitudes toward AI in art with self-determination theory as a theoretical framework, the general public's view of the integration of AI in the field of art is more negative compared to other fields in which AI was more easily adopted, such as mechanical engineering or real estate technology (Latikka et al. 2023)⁴. In fact, within art, the attitudes toward using AI in the subfield of art and culture were quite negative. The qualities that influence this perspective include relatedness and autonomy, where people are more accepting of the use of AI in art if they feel socially connected while using the tool and feel in control and free in their decision-making. Additionally, the study's results showed that those who regularly work with computers tend to hold more favorable views. Similarly, artists themselves and philosophers of art are highly against the use of AI in art, and generative AI: negative perceptions of AI creation of art are found to be linked with less favorable evaluations when participants believed the images were AI-generated (Hong & Curran 2019). Furthermore, artists argue that image generators are not artists and that while non-human entities can have aesthetic experiences and express emotion, a cultural product must be "controlled by a sensitivity to the attitude of the perceiver" to be considered a work of art (Jiang et al. 2023)⁵. In other words, art is the result of a process that is controlled for an end and isn't just the result of a spontaneous activity, such as image generation (Jiang et al. 2023)⁵. For artists, the human recognition of the intention of creation for an audience is what makes art itself; art is a form of communication. Artists perceive that image generators merely imitate the technical process and thus lack an understanding of the perspective of the audience or what the output is intended to communicate. Overall, both the general public and professional artists express a negative sentiment toward AI's role in art. This perspective highlights an urgent need for thoughtful analysis, weighing the complex benefits and drawbacks of AI integration as the art world navigates this transformative era.

Benefits

Despite the controversy surrounding the influx of image generators and the ethical disruption of art by such technology, algorithmic systems in art are also a great help to artists. For instance, animation is an artistic medium that is known for containing repetitive processes and being time-intensive. The work that animators do is extensive and very detailed, so they use many tools to reduce this arduous workload, such as

auto-rigging technology, which is enhanced by deep AI learning. AI can also automate tasks like creating backgrounds or crowds, or generating repetitive frames. In addition to reducing repetition, research has revealed that AI-assisted tools can help in 3D animation by positioning character bodies, lip-syncing characters using speech-aware software, or assisting character design by generating a variety of 3D designs and characters (Sharma & Juyal 2023)⁶. This support not only saves time and energy by speeding up the production process, but it also improves the quality of the animation by helping find references and leaving more time for animators to develop their original ideas (Sharma & Juyal 2023)⁶. In fact, in the production of Disney's *Frozen*, AI tools were used to add depth in the animation and sync facial expressions and sound. Research continues to show that AI has many more uses in color grading, sound output, and visual effects, and can even be employed in a feedback system during the rendering process.

Drawbacks

However, the use of AI tools in the field of art, even in animation, is more difficult than it seems. In animation, these benefits are acquired by extensive datasets, so these deep learning tools will be difficult for smaller businesses to use because of the substantial investment and technical expertise needed (Sharma & Juyal 2023)⁶. Also, in animation, due to simplifications of the frames (animations typically stick to simple depictions for enhanced emotional expression), generative AI may reinforce stereotypes because of the focus on fixed, simplified visual representations (Sharma & Juyal 2023)⁶. Additionally, image generators have also been seen to appropriate and distort identities of groups, encode biases, and reinforce stereotypes (Stark 2024)⁷. There are also many unanswered legal questions when it comes to determining how copyright law applies to image generators, and because many image generators are training using artists' works without their consent, digital artwork forgery has become a common reason for artists' mistrust.

Transparency and Trust

Trust and transparency are a must when considering opinions on AI in art, due to the legal considerations surrounding the use of artists' work for generative models. Indeed, a number of artists have "described this practice as 'invasive' and noted the manner in which it causes them reputational damage" (Jiang et al. 2023)⁵. Furthermore, in an observational survey, 80.17% of participants agreed that creators of AI systems should be required to provide detailed disclosure for what art they use to train the models and 61.87% believed that AI systems are a threat to those in art-related professions (Lovato et al. 2024)⁸. Additionally, a consensus is observed in many artists that "transparency and disclosure regarding the specific images used in training an AI model are essential pre-

requisites for them to deem the model fair and acceptable” (Lovato et al. 2024)⁸.

Anthropomorphism’s Role

Transparency is not the only leading factor in AI trust: anthropomorphism is one of the leading features in how people perceive fixed concepts like algorithmic systems, as “anthropomorphic features operate as social affordances [...] that simulate reciprocal engagements and foster a sense of trust between users and chatbots” (Maeda & Quan-Haase 2024)⁹. In animation, assigning human characteristics to mundane objects is common because it acts as a canvas that humans can project their feelings onto. On the integration of AI in art, the negative public perception could be attributed to the importance of anthropomorphism in the creative sector. It is noted that these public opinions relate to a lack of warm attributes and the threat to human creativity that automated AI systems pose (Latikka et al. 2023)⁴. Further, robot anthropomorphism and an increase in human attributes are seen to correlate with positive views on AI. A longitudinal study by Latikka et al. referencing a study by Bellaiche et al. exemplifies the importance of labeling: more positive appraisals were noted after the participants were aware of human engagement in the artistic process (Bellaiche et al. 2023)¹⁰. Latikka et al. also introduce the Self Determination Theory as a “theoretical framework for analyzing attitudes toward using AI in art from the basic human psychological needs perspective” (Latikka et al. 2023)⁴, where the psychological need of relatedness is defined as the need to experience connection with others. The study showed that greater perceived relatedness in the use of new technologies is connected to more positive attitudes toward using AI in art in general and in three art contexts (music, visual arts, and creating art). These findings suggested that the amount of social connection felt during the use of the technologies may correlate with positive user experiences, and that the feeling of control that individuals have over certain technology could explain the positive attitudes of AI. Similarly, additional findings from descriptive analyses demonstrate that AI in art may be perceived less positively than in other fields because of individuals’ tendency to view AI in artistic practice with concerns of the integration being “strange, false, cold, or even scary” (Latikka et al. 2023)⁴. Of course, many issues arise from the anthropomorphization of AI systems. Consequently, it is natural for the artists to argue that art is a uniquely human activity and cannot be initiated by an artifact. They claim that image generators have a fixed latent space, while human inspiration and relationship with their lived experiences change continuously and evolve over time (Latikka et al. 2023)⁴.

Threat vs. Augmentation

The significance of how using automation reduces artistic works’ emotional creativity is more powerful than one might initially think and may be a deal-breaker for many. Despite the

fear that the AI may replace human artists and that the products that AI produces may be of higher quality or productivity, some professionals argue that it is quite unlikely for AI to take animators’ jobs because “while AI can assist animators with specific tasks such as character rigging and motion graphics creation, it cannot replace animators’ imagination and artistic ability” (Sharma & Juyal 2023)⁶. This may be because creativity and uniqueness are often regarded as fundamental aspects of the art and culture field, and some have noted that “creativity is commonly seen as an exclusively human trait” (Latikka et al. 2023)⁴. Still, the integration of AI systems in the field of art through independent objects such as image generators has had many harmful connotations thus far. In fact, artists continue to face economic loss due to the influx of commercial image generators despite the previous argument. While comprehensive empirical data is still emerging, journalistic reports highlight immediate impacts; for instance, a Rest of World report cited by (Jiang et al. 2023)⁵ noted that a Chinese gaming industry recruiter has noticed a 70% drop in illustrator jobs, in part due to the widespread use of image generators and that another studio was reported to have laid off a third of its character design illustrators (Zhou 2023)¹¹. Furthermore, artists have even been approached by companies producing image generators, requesting their help in modifying the outputs of their systems.

Patterns

Both the public and artists themselves share a negative general sentiment; however, while negative public sentiment is geared towards the lack of social connection they feel, artists’ concerns range from concern about the human creativity in artificially created art to infringement of copyright law and data laundering. Also, only being able to see the inputs and outputs of models used in creative disciplines reflects a recurring trend where mistrust from both the public and the professionals is evinced. Most importantly for the specialty of art, the significance of emotional creativity in the field of art alone shows the increased gravity of the absence of robot anthropomorphism in most AI software used in art. In fact, research in this field indicates positive correlation of increased human engagement and perceived human values in the artistic processes with most positively inclined outside views.

Physical Healthcare & Medicine

Discrepancies, Benefits, and Drawbacks

In a survey of 600 US-based adults, around half of the survey participants (301 out of 600) reported a positive perception of AI’s future improvements in patient outcomes (Witkowski et al. 2024)¹². The majority of participants were more comfortable with AI being used for administrative tasks like scheduling patient appointments rather than those related to the doctor-patient relationship. On the other hand, there are

a number of concerns regarding the lack of humanization in AI, and mistrust of using such tools in medical procedures is popular amongst the public. While professions were seen to trust in medical algorithms, the general public was more prone to feelings of fear and discomfort (Witkowski et al. 2024)¹². However, a significant number of respondents also highlighted AI's potential to support medical professionals within the therapeutic alliance (a patient-clinician relationship). Correspondingly, in an online questionnaire, there was enthusiasm and optimism among these ear, nose, and throat (ENT) professionals on the integration of AI (Alshehri et al. 2024)¹³. The accuracy of AI-assisted diagnostics was generally seen as accurate, and they were overall satisfied with AI tools. Specifically, for those with 11–20 years of experience, 110 were satisfied out of the 152 professionals. In addition, the outlook on the future of AI systems in ENT is seen to be generally on the positive side, and satisfaction levels were slightly higher in ENT professionals working in hospital settings than in private practice. For the public, the sentiment of AI integration is very split, with many different concerns riddling the benefits of using AI in the direct medical sphere of the therapeutic alliance. However, professionals have a positive view on AI integration even within direct involvement in the healthcare field. Indeed, there are many benefits in the use of algorithmic systems in the medical domain, but drawbacks should be evaluated accordingly due to the significant role that healthcare plays in human lives.

Benefits

Some positive effects of the use of AI tools in healthcare include the predicted increase in patient autonomy through improving the accuracy of medical predictions. Previous research has also suggested that this integration may increase beneficence by providing reliable information and supporting proxy decision-making (Benzinger et al. 2023)¹⁴. Many researchers propose that the benefits of using AI-assisted tools provide additional support for clinical ethicists and reduce their moral burden in medical decision-making (Frantzén et al. 2011)¹⁵. Additionally, algorithmic systems in healthcare have been used successfully to optimize scheduling: simulation-based optimization (SBO) frameworks can be designed for real-time scheduling (Frantzén et al. 2011)¹⁵. Another type of algorithm used is genetic algorithms, a meta-heuristic optimizer that can generate optimal schedules. For example, a study on genetic algorithms' use in scheduling optimization introduces an algorithm that achieved up to 96% room utilization and outperformed the original scheduling by cutting patient wait times by up to two weeks (Rivera et al. 2020)¹⁶. Consequently, the implications for future use of optimization algorithms such as genetic algorithms may be generally positive due to their demonstrated ability to improve resource utilization in busy and complex environments.

Drawbacks

Conversely, a drawback of using algorithmic systems in medicine is that for some AI tools, there is an invasive nature that prevents their larger role. Scholars have noted a tendency of AI to show biases towards behaviors in the decision-making process, and have acknowledged that this could negatively affect minority populations (Benzinger et al. 2023)¹⁴. Further, the proposed benefit of algorithmic systems being supported for clinical ethicists could be undermined by arguments that the support could, in turn, increase stress and hurt professionals' confidence in their decisions when the AI does not agree with their assessment of the situation. Similarly, a view on how simplifying moral decisions to statistical correlations brings up concerns about lessening patient autonomy (Benzinger et al. 2023)¹⁴. On the other hand, the practicality of assimilating AI technology into healthcare systems like EMR systems is also an issue, because there exists compatibility challenges with diagnostic apparatus, which raises concerns on its usage and how this complexity may affect the reliability of diagnoses (Alshehri et al. 2024)¹³. Indeed, 218 of the respondents in the online questionnaire distributed to 600 certified and experienced ENT professionals highlighted sporadic diagnostic inaccuracies and a lack of strong validation across diverse clinical conditions. Another concern is economic and accessibility barriers, and 124 respondents were concerned about prohibitive costs that obstruct "accessibility in smaller clinical settings" and "availability of AI resources across various geographical locales and linguistic spectrums" (Alshehri et al. 2024)¹³. In general, implementing complex technology and software requires more technical assistance and system updates—two things that have been expressed as "suboptimal" by respondents (Alshehri et al. 2024)¹³. Overall, the range of the use of computational models in the medical field is still under construction, with the noticeable "lack of customization options for specific ENT cases", rigidity in clinical workflow adaptation, and apprehension about data privacy in algorithm-facilitated tools (Alshehri et al. 2024)¹³.

Transparency and Trust

Concerns such as privacy and cybersecurity are big contributors to patients' and professionals' trust. In medicine, trust is quite significant due to the focus on human beings' welfare. Consequently, it is natural that some are mistrustful of foreign concepts having a direct influence on humans' medical diagnoses. In fact, Witkowski et al.'s research on public perceptions of AI in healthcare stated that "73 people (12%) noted the fear/distrust they have for the use of AI in their care, indicating that they might be 'uncomfortable' with AI providing treatment and diagnosis independently of doctors" (Witkowski et al. 2024)¹². However, on the other hand, a majority of medical professionals expressed trust or neutrality in AI diagnostic results. They also generally trusted their level of confidence in explaining AI findings to patients: hospital

settings reported 70 and 76 participants expressing trust and full trust, respectively; private practices reported 50 respondents expressing trust; and academic institutions reported 30 respondents expressing trust (Alshehri et al. 2024)¹³. Still, the study noted that concerns regarding the understanding of AI systems posed as “barriers to trust and confidence” (Alshehri et al. 2024)¹³.

Anthropomorphism’s Role

Despite the field of medicine being one where the use of algorithmic systems is precedented (and its integration has much support), ultimately, AI’s lack of human characteristics still serves as an obstacle in its perceived capability to replicate the process of ethical deliberation, which may limit the effect of AI in this area. According to survey data from the State of Florida study of US-based adults, 62% of participants demonstrated concern about the lack of humanity of AI, which highlighted potential concerns about AI’s inability to provide that “human touch” that brings people together (Witkowski et al. 2024)¹². For instance, a participant in the study revealed, “I don’t think it is ideal to completely rely on computers, especially because a big part of healthcare is human interaction” (Witkowski et al. 2024)¹². Other participants referenced the lack of characteristics such as empathy, judgment, and subjective decision-making, which are considered crucial qualities in the “doctor-as-person element of patient-centered physical healthcare” (Witkowski et al. 2024)¹².

Threat vs. Augmentation

The public’s explicitly expressed worries about the limitations of most AI’s anthropomorphism are frequently introduced by the overall unease about the potential threat that AI poses to independent decision-making. For example, one participant from Witkowski et al.’s (2024)¹² previously mentioned study of U.S. survey data stated, “for appointments and reading scans, it wouldn’t bother me. I don’t want AI making my medical decisions.” Evidently, the unbreachable gap between AI and human decision-making causes apprehension in using AI in healthcare. Even if the patient has the final decision in their treatment, aversion to certain treatment methods because of the presence of AI could hinder the possible improvement and continued development of optimizing the use of algorithms in the medical field. The common fear of an overthrow of human principles in favor of fixed algorithms is one that seems to affect the public’s view on using automated methods for more significant uses in healthcare than simple tasks. Indeed, such concerns extend past the general fear of AI vs humans and include anxiety of AI’s replacement of humans in the workforce, the impact of AI in healthcare costs in the economy, and the reliability of the systems that algorithms produce. A takeaway from these results is the necessity of considering patient concerns and working to address them with mechanisms that guarantee transparency.

Patterns

In healthcare, while healthcare professionals share a positive sentiment towards the integration of algorithmic systems in medical settings, the public is without a general consensus due to several concerns. The reason for professionals’ outlook may be connected to their daily interaction with data and the algorithmic systems, unlike those from other industries. Medical professionals with more access to technical details report higher trust, while the public, with less understanding of the medical use of algorithms, may report fear or hesitation. For instance, Witkowski et al.’s survey yielded that respondents were least comfortable with AI administering medications and participating in doctors’ surgical procedures (Witkowski et al. 2024)¹².

Mental Healthcare

Discrepancies, Benefits, and Drawbacks

A cross-sectional survey found that overall, there is a low baseline knowledge of AI systems in mental healthcare, and that though some acknowledge benefits of using algorithmic systems, there is a substantial number of concerns and uneasiness surrounding this integration (Reading Turchioe et al. 2024)¹⁷. In addition, more participants with a history of mental illness than those without were uncomfortable with AI making a diagnosis of depression, but were comfortable sharing sensitive information with a mental health professional instead. They were also more concerned about AI making a wrong diagnosis, which highlighted the importance of transparency. Overall, other minorities in socio-demographics are more concerned about AI making a misdiagnosis, blame hospitals or clinics for errors, and have a safety expectancy with AI tools. For professionals, the perceived helpfulness of algorithmic tools is limited to the specific subfield, studies have found that “mental health professionals are not very positive towards AI performing treatment” (Sebri et al. 2020). The reason for the internal discrepancy may differ across subfields. For example, cognitive-behavioral psychotherapists have a positive view of the use of AI due to the potential for automating variables in mathematical computations that can be applied in cognitive-behavioral psychotherapy. In contrast, psychodynamic professionals may feel that a machine is not capable of interpreting the subjective meanings of unconscious states or relationships in their studies (Sebri et al. 2020). Negative opinions of the integration of algorithmic systems and AI in the field of mental health are prevalent among both the public and mental health professionals. Overall, there are causes of concern that are prevalent to both parties, and there is no significant discrepancy.

Benefits

A strength of AI systems in psychotherapy is providing objective information that can be found useful, yet not necessar-

ily ground-breaking. In a case study of ChatGPT's use in psychotherapy by Paolo Raile (2024)¹⁸, it was found that because of ChatGPT's robust understanding of theories of mind, "it is a good opportunity for people who already have psychotherapy experience to work alone with their material" (Raile 2024)¹⁸. Also, despite the lack of emotions that exist in AI, it may be refreshing for patients to freely interact with someone without social pressures or stigma.

Drawbacks

On the other hand, despite generative models initially showing empathy and flexibility in its suggestions, it was found to betray a strong bias in favor of Cognitive Behavioral Therapy (CBT) methods and fails to ask for more information (Raile 2024)¹⁸. The language model's tendency to employ a CBT approach to the situation could be considered as a case of breaching algorithmic ethics standards, and it may expose a liability for errors in the model. These one-sided outputs could be detrimental in the field of mental healthcare because the inherent bias could lead to ignorance of other therapeutic strategies (Raile 2024)¹⁸. In the case study, if the AI was asked about a specific approach, its answers were found to be broad and not profound enough to effectively help psychotherapists. Such answers showed a lack of diversity in psychotherapeutic approaches and a lack of relevance in the applications of the answer.

Transparency and Trust

Concerns about transparency were especially seen by vulnerable groups—those with a history of mental illnesses, those with poor mental health, ethnic minorities, older adults, women, and those with less financial resources—who expressed greater discomfort and fear about AI's integration (Reading Turchioe et al. 2024)¹⁷. These populations view AI as risky, require bioethical safeguards, and are prone to mistrust in providers if dangers from such tools appear. Within this study, more participants who rated their mental health as poor/fair were uncomfortable sharing sensitive information to help improve AI programs. They were also found to have a reliance on AI that impacted their increased mistrust in a mental health professional when the professional's assessment didn't align with AI's, and they were found to be more likely to be "somewhat" comfortable when receiving a diagnosis 98% accurate, despite being unexplainable, underscoring a smaller expectancy of transparency.

Anthropomorphism's Role

Having an innate human connection fosters empathy and understanding between the patient and the medical professional. However, when considering an AI as the proctor and not a fellow human, patients may not be as forthcoming with their personal data as they would with a human. Raile's study refers to another research paper by Carlbring et al. (2023)¹⁹,

which points out the limitations of a well-known example of an AI model: ChatGPT. Some of these limitations include ChatGPT and other models' lack of understanding of human metaphors such as irony, their limited ability to carry on long conversations (which may forestall therapeutic bonding), and their potentially unstable "accuracy, reliability, and jurisdictional impact" (Raile 2024)¹⁸. Many researchers have highlighted that mere imitations and programming appropriate responses may not be sufficient for those with unstable mental health (Carlbring et al. 2023)¹⁹. Carlbring et al. also recognize that communication between two entities is not limited to words, but goes beyond verbal conversation and includes body language and irony as well. When considering the difficulties of AI being able to interpret irony, the dangers of AI integration may outweigh the benefits—the paper states, "interpreting irony from a client is also a potential pitfall, which could lead to inappropriate therapeutic suggestions" (Carlbring et al. 2023)¹⁹.

Threat vs. Augmentation

Although algorithmic systems and AI models such as ChatGPT can be valuable, inexpensive, and offer a second outlook and coping strategies, researchers have made it clear that AI is best used in a supporting role rather than a main role in mental healthcare. For example, Raile (2024)¹⁸ communicated his strong belief that ChatGPT cannot replace a professional psychotherapist in his study, given the AI's curt and comparatively shallow insights. In responses regarding mental health, ChatGPT always "added that only a mental-health professional can make a diagnosis, and the questioner should consult one", actively referring the user to a human professional (Raile 2024)¹⁸.

Patterns

For professionals, the susceptibility to errors in AI levels of trust. For the public, the fear of professionals lacking an understanding of them as patients also reveals a fear of reliance. Because human understanding is necessary in jobs within this field and mental health is a direct factor in human lives, the need for human connection in patients makes it difficult for them to feel comfortable and may reduce the perceived threat. Moral stakes should be taken into consideration, as well as how subjectivity depends on different specialists: some AI systems may help in more objective jobs for computations for cognitive-behavioral psychotherapy, while in other areas, algorithmic systems may be insufficient in interpreting psychological meanings.

Education

Discrepancies, Benefits, and Drawbacks

Overall, there is a positive public response to the use of AI in education. Some students have a positive perception of

AI tutoring because the concept is thought to “improve teaching and learning into an adjustable learning environment and curriculum, and also promote personal learning experience” (Olaseni et al. 2024)²⁰. In another study, “94% of participants believe that an AI tool could improve their learning” (Latham & Goltz 2019)²¹. So, most participants thought they were useful and that if available, they would use them. However, 88% of participants still feel that humans play a significant role in education. On the other hand, the general perspective of teachers reports that 77% of participants have a negative inclination toward AI tutoring-based teaching (Olaseni et al. 2024)²⁰. This perspective can be attributed to multiple reasons: the counterproductive effects (as 43.7% stated), perceived incompetence (37.1%), anticipated replacement (9.8%), pedagogical beliefs (4.9%), and infrastructural deficits unsuitable for AI. Furthermore, it is important to note research that teachers who specialize in science and commercial (business units that deal with sales) departments are less likely to have a negative disposition toward AI tutoring compared to teachers who specialize in art departments (Olaseni et al. 2024)²⁰. It can be noted from the discrepancy that although many people recognize the vital role of human teaching, AI is readily accepted due to its simplicity. While such tools are commonly used by students already, this use may be frowned upon by professionals due to the importance of the process of learning and the role that human educators play in helping their students develop critical thinking and social skills.

Benefits

The prevalence of AI-assisted tools in learning has been widely promoted and independently used by many students. It is less time-consuming to use AI in studying because of reduction to travel time and flexibility in studying (Fitria 2021)²². Another advantage is AI’s ability to personalize a learning plan (e.g. Intelligent Tutoring Systems (ITS) individualize coaching) (Nikitina & Ishchenko 2024)²³. With this, students can have their weaknesses analyzed and obtain specific materials catered to their knowledge level and learning style, and virtual mentors can give comprehensive feedback (Fitria 2021)²². In a study where two groups each experienced 2 lessons: an AI-supported lesson or an active learning lesson, results showed that students learn more in less time when using the AI tutor and feel more engaged, and the authors considered the result to be due to AI’s self-paced aspect (Kestin et al. 2025)²⁴.

Drawbacks

However, it should be taken into consideration that AI tutoring may not be as influential where complex critical thinking is necessary for synthesizing multiple concepts. For example, findings show worse student performance when their discussions with AI “lacked structure and did not encourage critical thinking” (Kestin et al. 2025)²⁴. Finally, other core flaws in-

clude that system prompts are not very reliable, and AI algorithms tend to generate inaccurate hallucinations due to false confidence when providing an incorrect answer or marking replies (Kestin et al. 2025)²⁴. Further, social skills could be hindered due to less human interaction (Kestin et al. 2025)²⁴. Previous research has also stressed that if the learning process becomes very dependent on computing systems, software vulnerable to cyberattacks could affect students’ accessibility (Fitria 2021)²². Also, the relatively high cost to operate such systems could pose a problem for developing countries. Finally, with the prevalence of AI chatbots, unethical use of the tech could lead to new methods of plagiarism. Kestin et al.’s study presents concerns that unguided use leads to students not engaging in critical thinking and connects previous research findings to their study.

Transparency and Trust

The difference in a general positive public view and a more negative perception by professionals can be attributed to prior experience with AI. Because the general population generally has less direct experience with the inner workings of learning AI, they may start with more trust than those aware of the limitations and uses of such systems in education, as found in Ackermann et al.’s discussion (2025)²⁵. The major factor that education plays in “people’s subsequent employment and life opportunities” makes explainability, transparency, and accountability of AI vital (Vincent-Lancrin & Van der Vlies 2020)²⁶. In this field, trust is multi-dimensional: use of AI could also be untrustworthy if the information is accurate but misused, which has the potential to harm populations (Vincent-Lancrin & Van der Vlies 2020)²⁶. For example, a program could identify students at risk of dropping out to help them or to exclude them.

Anthropomorphism’s Role

Additionally, individuals with less experience and greater default trust may be more likely to view robots as living entities, whereas those with more experience can better distinguish between the two subjects. In learning, emotional connection is a minor factor in the effectiveness of AI tools that may be attributed to the objectiveness of the academic sector. Consequently, high anthropomorphism in pedagogical agents could be seen as out of place and even have negative effects on students’ comfort. Granted, previous research discusses how different communication styles can influence trust, and how “text-based communication can lead to dehumanization”, unlike more natural speech imbued with visible nonverbal cues, which might be helpful for learning (Ackermann et al. 2025)²⁵. However, in a study comparing two conditions—one in which a tutoring agent was presented only as a virtual avatar and another in which the agent was physically present in addition to the avatar—the findings showed that the physical presence of the agent had no significant ef-

fect on students' task enjoyment, perceptions of anthropomorphism, or task performance (Ackermann et al. 2025)²⁵. It was also noted that "students' progress on the task was lower when the robotic tutor agent was perceived as more sociable" or a disturbance due to certain human characteristics of the tutor (Ackermann et al. 2025)²⁵. These observations were recognized to possibly distract students from the current task and elicit negative reactions related to the "perception of danger and strangeness," such as discomfort (Ackermann et al. 2025)²⁵. This is supported by the phenomenon called the "Uncanny Valley", whereby artificially generated humanoids possess human-like features that evoke unease in the viewer. Therefore, anthropomorphism's role in the integration of AI in the field of education is not so significant in influencing people's opinions on this use unless there is high level—which may provoke a negative outlook due to unexpected humanization.

Threat vs. Augmentation

Because AI is typically not trained to follow constructive pedagogical practices, such as active learning, many researchers maintain that algorithmic systems should not replace in-person teaching but should instead allow students to achieve the most effective use of their time. Consequently, it is recommended for instructors to avoid using AI when students may misuse it to "circumvent critical thinking" (Kestin et al. 2025)²⁴. As technology advances, educators need to adapt to emerging learning devices so that they don't fall behind and risk being replaced. However, in the end, although AI can facilitate learning, it can not develop the character of the student the way teachers can, as supported by Fitria's (2021)²² observations. The irreplaceable role of the educators is not solely to teach the students, but it is to be a motivator and character builder for students through being a role model: teaching is about presenting topics to support students' learning (Nikitina & Ishchenko 2024)²³. Overall, most researchers do not see a threat of AI to human educators in the future, but the roles of teachers are certainly adapting: For instance, hybrid teaching of synchronous and asynchronous learning has been employed during the pandemic and is also a promising design for the future (Nikitina & Ishchenko 2024)²³.

Patterns

In the field of education, there is a discrepancy between both groups' opinions. First, not having much transparency in algorithmic models used in education is hypothesized to possibly lead to blind trust in the public, because of the little information provided to the public. Therefore, this trust may explain the public's positive view of the use of AI tutoring and education practices, while the more informed educators' views are negatively inclined. For students, algorithmic systems are often displayed in objective data-centered study tools, while educators' views on the role of algorithmic systems in their

work can be more precarious because of the values that teachers hold, and the role they play in students' lives is based on emotion as well, not just facts.

Discussion

As seen from the synthesis of the proposed factors on general opinion for each field, the intensity of these features provides detailed insight into the answer to this paper's research question. First, it was observed that there was a much greater role of robot anthropomorphism in fields other than that of education. Although AI's lack of human qualities played roles in all fields, the magnitude differed depending on the moral stakes in each sector (as noted in Claim 1) as well as the general objectivity or subjectivity in that field. For example, because medicine is an objective field, professionals may see the use of algorithmic systems as helpful to the advancement of the field, while the public may view their own health as a crucial interest that calls for safety and security. However, because this area also holds high moral stakes, this factor may precede a higher level of mistrust that influences public fear. An exception is the field of education, where although the public's use of AI in education is typically manifested in objective data-oriented study tools an educator's use of AI in their teaching is more frowned upon, due to the influential role they play in students' lives. As not only a tutor, but also a mentor and supporter, the subjectivity in which AI applies to the educator may differ. Research demonstrates how "perceptions of AI can be more positive regarding applications involving big data or other easily automated simple tasks compared to applications involving some aspect of human judgment and more complex tasks" (Latikka et al. 2023)⁴. Similarly, art is a very subjective and human-focused field, so the negative opinion of both the general public and professionals in the discipline may be explained by higher concerns of replacement and AI's lack of anthropomorphism. These notions match contrasting findings from other sources on people's attitudes toward AI in applications involving large sets of data and easily automated tasks compared to others that involve more subjectivity and complexity, such as art, which warrants more human judgment. Furthermore, it was noticed that mistrust and fear were more apparent in fields such as the healthcare and mental healthcare domains, which hold high moral weight and stakes for humans and their direct well-being and lives. In areas with high connotations of human qualities (such as the area of art), perceived threat and negative inclinations are present when there is a lack of anthropomorphism. Trust is also a factor that can be swayed, with AI's predisposed decisions being another influence. Trust in the general public and the field's professionals differed in the fields of education and healthcare. In the case of the discrepancies between the trust of the public and professionals in these areas of practice, an explanation may be that

a lack of information in the public can also lead to blind trust (in the case of education) rather than just mistrust (as seen in the other 3 fields).

As shown in Table 1 below, perceptions of AI differed consistently according to the degree of subjectivity and moral consequence associated with each field, which influenced similar factors such as trust and perceived threat or a preference for AI assistance. Domains involving emotional interpretation, creativity, or interpersonal trust seemed to evoke stronger concerns regarding anthropomorphism and perceived threat, while more objective fields and/or data-oriented applications of AI were perceived more favorably when AI functioned in an assistive role. In addition, the factors of transparency and trust emerged as recurring determinants of acceptance, especially from public attitudes. Overall, the cross-study comparison reveals several recurring similarities across the reviewed fields.

The general trends noticed in this literature review were that the higher the stakes of the field, the more important the role of robot anthropomorphism, which amplified distrust and perceived threat of AI and algorithmic systems in the field. Transparency in the deep learning models was also researched thoroughly, as it is hypothesized that less transparency negatively affects trust, as a lack of transparency leads to a lack of information and understanding, which in turn leads to distrust in most cases (there may be an effect of increased trust depending on the amount of information made available to the public). Finally, it is important to note the perceived threat (which may be affected by mistrust and lack of transparency and negatively correlated with perceived anthropomorphism). Altogether, these findings suggest that attitudes towards AI are impacted not just by the technological capabilities of a system, but also by the social and environmental aspects of the domain. This analysis supports the paper's argument that perceptions of AI are closely connected to subjective judgment and moral responsibility.

Ethical Interpretation of Findings

The findings of this review may also be interpreted through established ethical frameworks. In more objective and efficiency-oriented applications, such as scheduling optimization or technical assistance, support for AI systems tends to reflect utilitarian perspectives that prioritize improved outcomes and efficiency. However, resistance in the other three areas more closely aligns with deontological ethical frameworks, which place more emphasis on moral responsibility, emotion, and the intrinsic value of human judgment. It can be seen that in these fields, participants may view transparency, accountability, and anthropomorphism as ethical requirements that AI systems must fulfill. Additionally, many of the concerns regarding trust and transparency that arose in these fields align

with ethical frameworks that emphasize moral values and advocate for human-centered AI integration.

Limitations

This review has several limitations that should be acknowledged. First, the paper synthesizes highly distinct fields, which differ substantially in their institutional structures. As a result, broad cross-disciplinary comparisons may oversimplify important contextual differences between domains. The analyzed public perceptions of AI may also vary significantly across cultures, technological environments and scope, and institutional settings, which were factors that were not covered in this review. Additionally, the studies synthesized in this paper examine a wide range of algorithmic systems with differing functions and levels of autonomy, including diagnostic systems, generative models, recommendation systems, conversational agents, optimization algorithms, and unofficial tutoring systems. Although these technologies were grouped comprehensively under the category of “algorithmic models” to identify overarching patterns in public and professional perceptions, these distinctions may influence the analyzed factors differently. Another limitation is that some arguments within the review rely more heavily on qualitative interpretation supported by theoretical evidence rather than on direct quantitative comparison across studies. Although the trends and recurring themes were arranged systematically, the selected studies differ in methodology and sample populations. Furthermore, certain sections like that of mental healthcare, draw partly from the exploratory evidence rather than controlled clinical studies. Studies such as these should not be interpreted as definitive evidence of psychotherapeutic competence. Consequently, these findings should be weighted accordingly. The use of narrative literature review as the primary research method also introduces potential limitations regarding the scope of the source selection. The reviewed studies were selected from accessible academic databases and English-language publications, which may exclude relevant perspectives from other databases. Many studies focused on populations from the United States or Europe as its geographical range, with exceptions being occasional contextual evidence from studies based in Asia and Nigeria. This established scope is helpful to narrowing the results analyzed in the review. However, future research would benefit from broader cross-cultural comparisons and greater representation of underexamined populations and technical environments. Finally, the sources compiled by the paper did not thoroughly examine long-term real-world deployment of AI systems. Consequently, participants' attitudes may not fully reflect future behaviors or perceptions in the domains where AI integration becomes more widespread and/or normalized.

Table 1 Cross-Study Comparative Analysis of AI Perceptions Across Fields

Field	Primary AI Function	Subjectivity	Moral Stakes	Public Attitude	Professionals' Attitudes	Main Trust Concern	Anthropomorphism	Perception
Art	Generative creation	Highly subjective	Moderate	Generally cautious	Strongly negative	High: Ownership, authenticity	High	Threat
Physical Healthcare	Scheduling, Optimization	Objective	Very high	Mixed due to safety concerns	Mixed; accepting of usage	Moderate: Transparency, errors	Moderate	Augmentation
Mental Healthcare	Conversational assistance	Subjective	Very high	Generally distrustful	Skeptical	High: Lack of empathy, misdiagnostics	Very high	Augmentation
Education	Personalized learning	Mixed	Moderate	Generally positive	Skeptical	Low: Overreliance, impact on students	Low	Augmentation

Conclusion

In conclusion, the findings of this literature review suggest that the influence of human judgment and algorithmic models varies across fields depending on the degree of subjectivity and human-focused stakes involved. The review supports the hypothesis that moral and emotional reasoning shape public and professionals’ attitudes toward AI. The significance of researching this is to demonstrate the human impact of advancing in technical sectors with AI development instead of solely focusing on the statistical changes in our world. Therefore, the research on this topic was investigated with hopes of shedding light on its negatives and positives, multiple perspectives on the integration of AI in different fields, and elements that influence opinion in order to show agency for improving the social impact of such advancements. The conclusion of the context that impacts professional and public perceptions of algorithmic authority may be very helpful to further studies. It is important for future research to focus on examining how these mediating factors can be seamlessly incorporated into AI use, initiating ethical accountability for both the public and field professionals. The results also suggest that AI integration should prioritize collaboration between humans and AI systems rather than full automation, especially in high-stakes fields. For instance, in healthcare settings, policymakers should implement systems primarily as assistive tools that support professional clinicians and also require transparency regarding how and why certain algorithmic recommendations were generated. Setting explainability standards and adjusting training for clinicians may help strengthen public trust, calm unease in patients, and reduce fear and risk in medical contexts. Similarly, in educational contexts, institutions should continue to gauge the effects of AI integration into school curriculum and evaluate biases of such tools accordingly. It may be beneficial for educators to receive guidance on balancing AI-assisted learning with the preservation of mentorship and creativity. Overall, for policymakers, it is important to distinguish between assistive, generative, and diagnostic technologies and their independent reputation in order to stay aligned with the integration of AI. It is also recommended for further studies to examine longitudinal changes in public and professional attitudes toward AI as exposure increases, as well as demographic variations in emotional responses to algorithmic models. Some direct proposed future steps in the real world to overcome the barrier of distrust include implementing hybrid

approaches and ensuring responsible AI adoption through the enhancement of public literacy. For instance, for education, strategies that consider the long-term impact of AI on students may have favorable outcomes (Olaseni et al. 2024)²⁰. AI literacy in clinicians is also recommended to be improved, as well as adaptations of “AI to patient-centered care ... along with specific guidance for doctors on implementing AI within different medical settings” (Witkowski et al. 2024)¹². It has also been stressed that the user-centered design of AI should “take into account professionals’ needs and representations” and “the characteristics of specific fields and approaches” (Sebri et al. 2020). Finally, as seen in mental healthcare, where it has been noted that “openly sharing personal experiences as a means of enhancing the therapeutic alliance and improving depression treatment outcomes”, the application of anthropomorphism as needed could be an effective course of action (Carlbring et al. 2023)¹⁹. In fact, research points out that positive and fulfilling experiences with technology can help reduce resistance to AI in creative fields (Latikka et al. 2023)⁴. In addition, interpretability in AI behaviors would also increase people’s acceptance of AI. Therefore, possible steps can also include encouraging collaboration between fields and the alignment of tech innovation with human values. Ultimately, developing ethical AI frameworks that employ ethical transparency, accountability, machine learning fairness, emotional awareness, and user-centered design could advance the responsible coexistence of human and algorithmic judgment in our modern society.

Acknowledgement

Thank you for the guidance of Ellie Fassman from Cornell University in the development of this research paper.

References

1 International Organization for Standardization, *What is Artificial Intelligence (AI)?*, 2024, <https://www.iso.org/artificial-intelligence/what-is-ai>.

2 C. Roselli, L. Lapomarda and E. Datteri, *How culture modulates anthropomorphism in human-robot interaction: A Review*, 2025, 10.1016/j.actpsy.2025.104871.

3 Y. Zhang, J. Wu, F. Yu and L. Xu, *Moral judgments of human vs. AI agents in moral dilemmas*, 2023, 10.3390/bs13020181.

4 R. Latikka, J. Bergdahl, N. Savela and A. Oksanen, *AI as an artist? A two-wave survey study on attitudes toward using artificial intelligence in art*, 2023, 10.1016/j.poetic.2023.101839.

- 5 H. H. Jiang, L. Brown, J. Cheng, M. Khan, A. Gupta, D. Workman, A. Hanna, J. Flowers and T. Gebru, *AI art and its impact on artists*, 2023, 10.1145/3600211.3604681.
- 6 H. Sharma and A. Juyal, *Future of animation with Artificial Intelligence*, 2023, 10.29121/shodhkosh.v4.i2se.2023.559.
- 7 L. Stark, *Animation and artificial intelligence*, 2024, 10.1145/3630106.3658995.
- 8 J. Lovato, J. W. Zimmerman, I. Smith, P. Dodds and J. L. Karson, *Foregrounding artist opinions: A survey study on transparency, ownership, and fairness in AI generative art*, 2025.
- 9 T. Maeda and A. Quan-Haase, *When human-AI interactions become parasocial: Agency and anthropomorphism in affective design*, 2024, 10.1145/3630106.3658956.
- 10 L. Bellaiche, R. Shahi, M. H. Turpin, A. Ragnhildstveit, S. Sprockett, N. Barr, A. Christensen and P. Seli, *Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork*, 2023, 10.1186/s41235-023-00499-6.
- 11 V. Zhou, *AI is already taking video game illustrators' jobs in China*, Rest of World, 2023, <https://restofworld.org/2023/ai-image-china-video-game-layoffs/>.
- 12 K. Witkowski, R. B. Dougherty and S. R. Neely, *Public perceptions of artificial intelligence in Healthcare: Ethical concerns and opportunities for patient-centered care*, 2024, 10.1186/s12910-024-01066-4.
- 13 S. Alshehri, K. A. Alahmari and A. Alasiry, *A Comprehensive Evaluation of AI-Assisted Diagnostic Tools in ENT Medicine: Insights and Perspectives from Healthcare Professionals*, 2024.
- 14 L. Benzinger, F. Ursin, W.-T. Balke, T. Kacprowski and S. Salloch, *Should artificial intelligence be used to support clinical ethical decision-making? A systematic review of reasons*, 2023, 10.1186/s12910-023-00929-6.
- 15 M. Frantzén, A. H. C. Ng and P. Moore, *A simulation-based scheduling system for real-time optimization and decision making support*, 2011, 10.1016/j.rcim.2010.12.006.
- 16 G. Rivera, L. Cisneros, P. Sánchez-Solís, N. Rangel-Valdez and J. Rodas-Osollo, *Genetic algorithm for scheduling optimization considering heterogeneous containers: A real-world case study*, 2020, 10.3390/axioms9010027.
- 17 M. Reading Turchioe, P. Desai, S. Harkins, J. Kim, S. Kumar, Y. Zhang, R. Joly, J. Pathak, A. Hermann and N. Benda, *Differing perspectives on artificial intelligence in mental healthcare among patients: A cross-sectional survey study*, 2024, 10.3389/fdgth.2024.1410758.
- 18 P. Raile, *The usefulness of ChatGPT for psychotherapists and patients*, 2024, 10.1057/s41599-023-02567-0.
- 19 P. Carlbring, H. Hadjistavropoulos, A. Kleiboer and G. Andersson, *A new era in internet interventions: The advent of chat-GPT and AI-Assisted therapist guidance*, 2023, 10.1016/j.invent.2023.100621.
- 20 V. M. Olaseni, *Teachers' perception towards integration of artificial intelligence tutoring-based system in the school curriculum: A survey*, 2024, 10.38159/ehass.202451319.
- 21 A. Latham and S. Goltz, *A survey of the general public's views on the ethics of using AI in Education*, 2019, 10.1007/978-3-030-23204-7_17.
- 22 T. N. Fitria, *Artificial Intelligence (AI) In Education: Using AI Tools for Teaching and Learning Process*, 2021.
- 23 I. Nikitina and T. Ishchenko, *The impact of AI on teachers: Support or replacement?*, 2024, 10.23856/6511.
- 24 G. Kestin, K. Miller and A. e. a. Klaes, *AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting*, 2025, 10.1038/s41598-025-97652-6.
- 25 H. Ackermann, A. Henke, J. Chevalère, H. S. Yun, V. V. Hafner, N. Pinkwart and R. Lazarides, *Physical embodiment and anthropomorphism of AI tutors and their role in student enjoyment and performance*, 2025, 10.1038/s41539-024-00293-z.
- 26 S. Vincent-Lancrin and R. van der Vlies, *Trustworthy Artificial Intelligence (AI) in Education*, 2020, 10.1787/a6c90fa9-en.

Appendix: Literature Review Matrix

Table 1A. Methodological Classification of Empirical Studies

Study	Intervention Type	User Group	Decision Task	Autonomy	Outcome Measured	Evidence Type
Zhang et al. (2023) ³	Prediction/Decision making	General public	Moral judgments of humans versus AI	Autonomous	Blame and trust attribution	Empirical study
Latikka et al. (2023) ⁴	Generation	General public	Attitudes toward using AI in different contexts	Assistive	AI autonomy effects	Empirical survey data
Lovato et al. (2024) ⁸	Generation	Adult participants	Attitudes of generative AI in art	Assistive	Evaluation of conflicts from AI	Empirical survey data
Alshehri et al. (2024) ¹³	Diagnosis	ENT professionals	Adoption of AI diagnostic tools	Decision support	Satisfaction, confidence	Empirical survey data
Rivera et al. (2020) ¹⁶	Optimization	Hospital administrators	Scheduling optimization	Semi-autonomous	Efficiency	Empirical data
Witkowski et al. (2024) ¹²	Prediction	U.S.-based adults	AI involvement in medical care	Decision support	Comfort, trust	Empirical survey data
Reading Turchioe et al. (2024) ¹⁷	Diagnosis, Recommendation	Online sample of adults	AI mental health diagnosis	Decision support	Comfort, trust, blame	Hypothetical scenarios
Raile (2024) ¹⁸	Conversational assistance	Simulated psychotherapists and patients	ChatGPT therapeutic interaction	Assistive	Adequacy, helpfulness, bias	Simulated interactions
Latham & Goltz (2019) ²¹	Recommendation	Adult learners	AI-supported education scenarios	Assistive	Usefulness, acceptance	Survey data/Hypothetical scenarios
Ackermann et al. (2025) ²⁵	Conversational assistance	Secondary school students	Interaction with an intelligent tutoring system	Assistive/Interactive	Engagement, performance	Simulated learning environment
Kestin et al. (2025) ²⁴	Recommendation	College students	AI-powered tutor vs active learning class	Assistive	Learning efficiency	Empirical data