# Psychological Impact of Artificial Intelligence on Human Cognition, Behavior, and Emotional Well-Being: A Systematic Review

**Reuben B Joy**

**Psychologist**

**DOI: https://doi.org/10.47772/IJRISS.2026.10100046**

**Received: 10 July 2026; Accepted: 15 July 2026; Published: 28 July 2026**

## ABSTRACT

The rapid proliferation of artificial intelligence across diverse contexts has generated increasing scholarly attention to its psychological consequences. Despite growing empirical literature, evidence on AI's cognitive, behavioral, and emotional impacts remains fragmented across disciplinary silos, precluding an integrated understanding. This systematic review synthesized evidence across cognition, behavior, and emotional well-being, identifying mechanisms shaping long-term human functioning. Following PRISMA 2020 guidelines, six databases (PsycINFO, PubMed, Scopus, ERIC, Google Scholar, Web of Science) were searched. Out of the 450 identified records, 57 peer-reviewed studies met inclusion criteria and were coded across 28 cognitive, behavioral, and emotional constructs. Thematic synthesis yielded three major clusters. Cluster A (n = 18): cognitive offloading, critical thinking erosion, and metacognitive atrophy were dominant, moderated by contextual and pedagogical factors. Cluster B (n = 19): AI drove adoption and learning gains but eroded agency, fostered algorithmic dependency, and produced context-sensitive trust calibration. Cluster C (n = 20): therapeutic AI reduced depressive symptoms (small-moderate effects), whereas social media AI amplified anxiety and loneliness, and companionship AI increased isolation over 12 months despite short-term comfort. AI functions as a cognitive enhancer, behavioral modifier, emotional support, and dependency mechanism, with effects moderated by design, context, user literacy, and exposure duration. Findings call for AI literacy, human-centered design, and evidence-based policy to ensure AI strengthens rather than supplants human psychological capacity.

**Keywords:** Artificial Intelligence, Cognitive Offloading, Algorithmic Trust, Emotional Well-Being, Human Agency, AI Literacy, Systematic Review, Mental Health

## INTRODUCTION

### Background: The Pervasive Rise of Artificial Intelligence

Artificial intelligence has undergone a transformation from a specialist computational domain into a pervasive dimension of everyday human experience. Within fewer than two decades, AI systems have penetrated education, healthcare, professional practice, interpersonal communication, and domestic life at a pace unprecedented in the history of technological development. From intelligent tutoring systems and adaptive learning platforms that personalize educational content to algorithmic recommendation engines that curate information, entertainment, and social contact, from conversational AI agents that deliver therapeutic support to autonomous decision-support systems that assist in clinical diagnosis and financial planning- the scale and intimacy of AI's involvement in human life are without historical precedent (Luckin, 2018; Russell, 2019). This expansion has been particularly dramatic in the aftermath of large language models (LLMs) such as GPT-3 and its successors. The release of ChatGPT in late 2022 represented an inflection point, placing generative AI capabilities directly in the hands of hundreds of millions of users worldwide and fundamentally altering the cognitive, communicative, and relational landscape within which human beings operate. As Dwivedi et al. (2023) observed in their synthesis of 43 expert perspectives, generative AI's capacity to produce human-quality text, images, and analytical reasoning at minimal cost raises questions not merely about productivity or economic

disruption, but about the foundational psychological processes through which humans learn, decide, relate, and regulate their emotional lives.

The integration of AI into educational environments illustrates the scope of this transformation. AI-based tutoring systems now provide adaptive instruction to millions of learners across primary, secondary, and tertiary educational settings (Woolf, 2010; Zawacki-Richter et al., 2019). In professional contexts, AI decision-support tools are reshaping the nature of expert judgment in medicine, law, finance, and engineering (Dwivedi et al., 2021). In the domain of mental health, conversational AI chatbots are delivering evidence-based therapeutic content to populations who would otherwise lack access to care (Fitzpatrick et al., 2017; Abd-Alrazaq et al., 2020). And in the social sphere, AI-driven platform algorithms govern the content, sequence, and emotional valence of social information that billions of users encounter daily (Montag & Hegelich, 2020; Meshi & Ellithorpe, 2021). The psychological stakes of these developments are substantial.

**Psychological Relevance: Why Cognition, Behavior, and Emotion Must Be Studied Together**

The psychological impact of AI does not resolve neatly into any single domain of psychological science. Cognitive processes, including attention, memory, critical thinking, decision-making, and metacognition, are closely linked to behavioral patterns. These behaviors influence how individuals interact with technology, acquire knowledge, make decisions, and regulate their daily habits. At the same time, cognitive and behavioral processes both affect and are affected by emotional states, including anxiety, well-being, loneliness, and self-efficacy. An individual who relies on AI for decision support does not merely exhibit a cognitive shift toward automation; the behavioral habit of deference to AI reinforces itself through emotional relief from the anxiety of autonomous choice, which in turn shapes subsequent cognitive strategies. These interactions are dynamic, reciprocal, and temporal.

Yet the existing literature has largely studied these domains in isolation. Cognitive scientists have examined AI's effects on memory and attention (Sparrow et al., 2011); educational technologists have measured learning behavior and engagement (Woolf, 2010); clinical psychologists have evaluated AI chatbots for depression and anxiety (Abd-Alrazaq et al., 2020); social scientists have theorized AI's relational consequences (Turkle, 2011). The result is a body of evidence that is rich in its individual contributions but fragmented in its collective architecture, a limitation that has prevented the emergence of an integrated psychological understanding of how AI reshapes the human mind, the human behavioral repertoire, and the human emotional life across their points of intersection.

**Research Gap**

Several specific gaps motivate the present review. First, no prior systematic review has simultaneously synthesized evidence across cognitive, behavioral, and emotional dimensions of AI's psychological impact using a unified, cross-domain coding framework. Existing reviews have focused on specific applications (e.g., educational AI, health chatbots) or specific constructs (e.g., algorithmic trust, AI-driven depression reduction) rather than on the integrated psychological system through which AI's effects propagate. Second, the interaction effects between cognitive, behavioral, and emotional outcomes, which are fundamental to understanding long-term human functioning in AI environments have not been systematically mapped. Third, the growing diversity of AI modalities (generative LLMs, ITS, social robots, algorithmic recommendation systems, voice assistants) demands a synthesis that can identify both general psychological mechanisms and modality-specific effects. Fourth, the moderating roles of contextual factors such as developmental stage, AI literacy, exposure duration, task complexity, and structural equity, all require integrated examination.

**Research Objectives**

The primary objective of this systematic review was to synthesize peer-reviewed evidence on the psychological impact of AI across cognitive, behavioral, and emotional domains, and to identify integrated mechanisms through which AI shapes long-term human psychological functioning. Secondary objectives included:

a) identifying the most frequently documented psychological constructs associated with AI interaction
b) examining the directional valence (positive, negative, or mixed) of AI's effects across each domain
c) analyzing contextual and individual factors that moderate the direction and magnitude of psychological effects, and
d) developing an evidence-based integrated conceptual model of AI's psychological impact.

### Research Questions

1) How does AI influence human cognitive functioning, including memory, attention, critical thinking, decision-making, cognitive offloading, creativity, problem-solving, and metacognition?
2) How does AI influence human behavioral patterns, including technology adoption, learning behavior, engagement, motivation, productivity, trust, agency, and dependency?
3) How does AI affect human emotional well-being, including anxiety, stress, loneliness, emotional support, self-efficacy, social comparison, emotional dependency, human relationships, and mental health?
4) What integrated psychological mechanisms emerge from the joint examination of cognitive, behavioral, and emotional evidence, and what do these mechanisms imply for long-term human functioning in AI-saturated environments?

## METHODOLOGY

### Review Design

This study employed a PRISMA 2020-compliant systematic review with thematic synthesis as its primary analytical strategy. Thematic synthesis was selected over meta-analysis given the heterogeneity of study designs, AI modalities, and psychological outcomes across the included literature. The review protocol was structured around a predetermined three-domain coding framework (cognitive, behavioral, emotional) to enable cross-thematic integration.

### Databases Used

Systematic searches were conducted across six major academic databases: PsycINFO, PubMed, Scopus, ERIC (Education Resources Information Center), Google Scholar and Web of Science. These databases were selected to ensure coverage across psychological, clinical, educational, and interdisciplinary AI research.

### Search Strategy

Searches were conducted using combinations of the following keywords and Boolean operators:

("artificial intelligence" OR "AI" OR "machine learning" OR "large language model" OR "ChatGPT" OR "intelligent tutoring system" OR "conversational agent" OR "social robot") AND ("psychology" OR "cognition" OR "cognitive" OR "memory" OR "attention" OR "critical thinking" OR "decision-making" OR "behaviour" OR "learning" OR "engagement" OR "motivation" OR "emotion" OR "well-being" OR "anxiety" OR "depression" OR "loneliness" OR "mental health" OR "self-efficacy")

Search terms were adapted to the controlled vocabulary of each database (e.g., MeSH terms in PubMed). Reference lists of identified systematic reviews were hand-searched for additional relevant studies.

### Inclusion Criteria

- Peer-reviewed empirical studies, systematic reviews, meta-analyses, or theoretically grounded conceptual analyses.
- Published in English between 2010 and 2026, capturing the contemporary AI landscape including LLM emergence.

- Examining psychological outcomes (cognitive, behavioral, or emotional) in relation to AI interaction or AI system deployment.
- Involving human participants or reporting human-centered psychological data.
- Published in indexed academic journals or conference proceedings with peer review.

**Exclusion Criteria**

- Studies examining exclusively technical AI development without psychological outcomes.
- Grey literature, dissertations, or non-peer-reviewed reports.
- Studies not available in English.
- Duplicate publications of the same data.
- Studies published before 2010 (except seminal theoretical works cited in the included literature).

**Screening Procedure**

Screening proceeded in two stages. At Stage 1, titles and abstracts of all 380 deduplicated records were screened against inclusion and exclusion criteria. At Stage 2, 150 full-text articles were retrieved and assessed for eligibility. Reasons for exclusion at full-text stage were recorded. Discrepancies were resolved through discussion and reference to the protocol.

**PRISMA Flow Description**

An initial database search across six sources yielded 420 records, supplemented by 30 additional records identified through reference list searching, for a total of 450 records. Following deduplication, 380 unique records were screened at the title and abstract level; 230 records were excluded for failing to meet the inclusion criteria. Full-text assessment was conducted on 150 articles; 93 were excluded with reasons (predominantly: no psychological outcome data, technical AI papers without human participant data, or failure to meet peer-review criteria). A final sample of 57 studies was included in the review, distributed across three themes: Theme A - Cognitive Impact (n = 18), Theme B - Behavioral Impact (n = 19), Theme C – Emotional Impact (n = 20) (see Figure 1 for the PRISMA flow diagram).

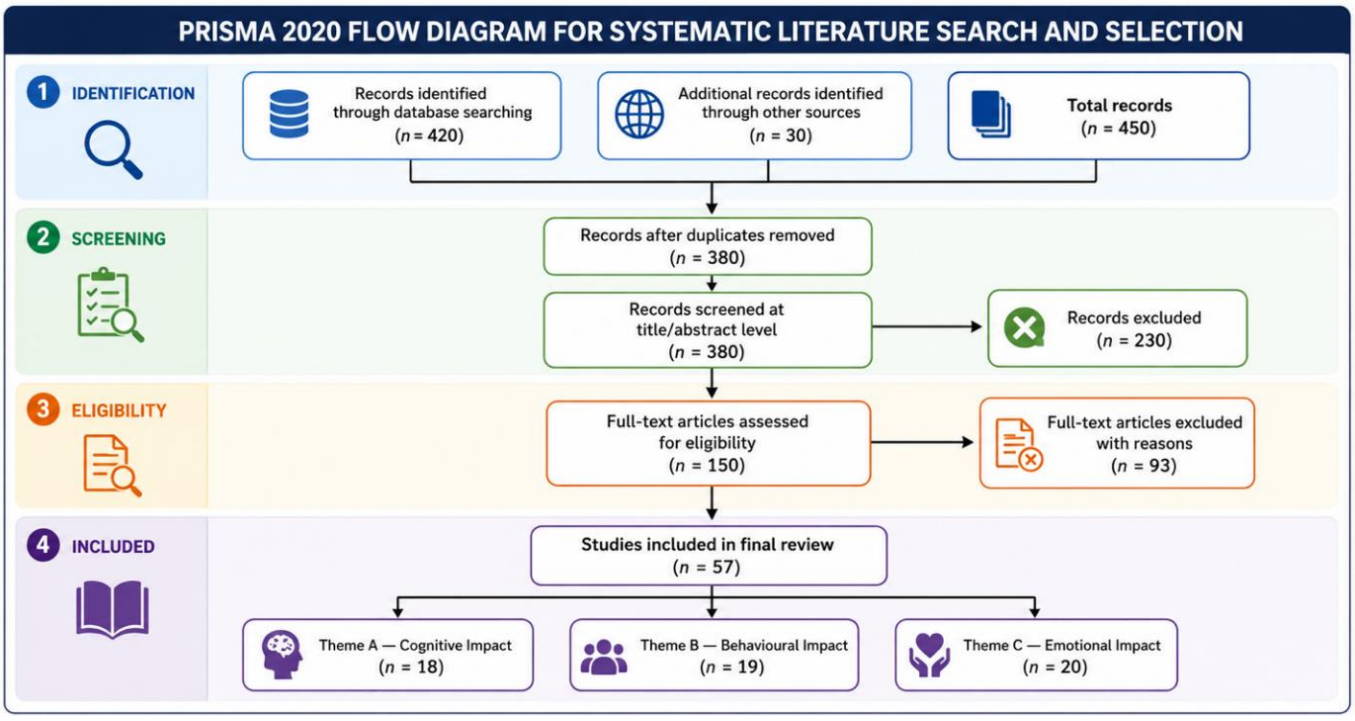


Figure 1: PRISMA Flow Diagram

**Data Extraction**

Standardized data extraction covered: author(s) and year, country, study design, sample size, AI context or tool, psychological domain(s) addressed, key findings, methodological limitations, and evidence quality classification. The methodological quality of the included studies was assessed using the Joanna Briggs Institute (JBI) Critical Appraisal Checklists appropriate to each study design (Joanna Briggs Institute, 2020), following the guidance outlined in the JBI Manual for Evidence Synthesis (Aromataris et al., 2024). The checklist was selected according to the study design (e.g., quasi-experimental, analytical cross-sectional, cohort, randomized controlled trial, qualitative, or systematic review). Each study was evaluated against the relevant criteria relating to sampling, measurement, data analysis, and risk of bias. Sources that were conceptual, theoretical, or opinion-based without primary empirical data were classified as not applicable for formal appraisal, consistent with established JBI guidance. The appraisal informed the interpretation of findings during synthesis, with greater emphasis placed on evidence from studies demonstrating stronger methodological rigor.

**Coding Framework**

Extracted data were coded against a structured framework of 28 psychological constructs spanning three domains (Table 1). Constructs were derived inductively from the literature during initial screening and refined iteratively. Multiple codes per study were permitted where evidence addressed more than one construct.

Table 1: Coding Framework

| Theme | Code | Construct | Definition |
|---|---|---|---|
| **COG** | **MEM** | Memory | Effects on encoding, retention, and recall of information |
| **COG** | **ATT** | Attention | Changes in sustained attention and attentional allocation |
| **COG** | **CT** | Critical Thinking | Capacity to evaluate, analyse, and synthesise information |
| **COG** | **DM** | Decision-Making | Autonomous versus AI-assisted decision processes |
| **COG** | **CO** | Cognitive Offloading | Externalisation of cognitive tasks to AI systems |
| **COG** | **CRE** | Creativity | Generative, divergent, and transformational ideation |
| **COG** | **PS** | Problem Solving | Application of reasoning to novel or structured challenges |
| **COG** | **META** | Metacognition | Self-monitoring and regulation of cognitive processes |
| **BEH** | **DEP** | Dependency | Over-reliance on AI at the expense of independent functioning |
| **BEH** | **LB** | Learning Behaviour | Changes in study habits, knowledge acquisition, and retention |
| **BEH** | **ENG** | Engagement | Active participation and immersion in AI-mediated tasks |
| **BEH** | **PROD** | Productivity | Efficiency and output quality in AI-assisted work |
| **BEH** | **MOT** | Motivation | Intrinsic and extrinsic drive in AI contexts |
| **BEH** | **DB** | Decision Behaviour | Overt choice behaviour in AI advisory situations |
| **BEH** | **ADOPT** | Technology Adoption | Decision to use and continue using AI tools |
| **BEH** | **TRUST** | Algorithmic Trust | Confidence calibration in AI systems versus humans |
| **BEH** | **AGENCY** | Human Agency | Preservation or erosion of autonomous choice and action |
| **BEH** | **BC** | Behavioural Change | Observable changes in health or lifestyle behaviour |
| **EMO** | **ANX** | Anxiety | State or trait anxiety associated with AI interaction |
| **EMO** | **STR** | Stress | Perceived stress arising from or reduced by AI use |
| **EMO** | **LON** | Loneliness | Social isolation in AI-mediated relational environments |
| **EMO** | **ES** | Emotional Support | Perceived empathy, validation, and care from AI systems |
| **EMO** | **SE** | Self-Efficacy | Belief in one's capacity to cope and perform |
| **EMO** | **WB** | Well-Being | Hedonic and eudaimonic psychological well-being |
| **EMO** | **SC** | Social Comparison | Evaluative comparisons facilitated by AI platforms |
| **EMO** | **ED** | Emotional Dependency | Reliance on AI for emotional regulation and connection |

| | | | |
|---|---|---|---|
| **EMO** | **HR** | Human Relationships | Quality and depth of interpersonal bonds |
| **EMO** | **MH** | Mental Health | Depression, anxiety disorders, and psychological adjustment |

### Synthesis Approach

Thematic synthesis was conducted in three stages following Thomas and Harden's (2008) framework:

a) line-by-line coding of key findings from included studies against the 28-construct framework;
b) development of descriptive themes capturing the direction, frequency, and consistency of evidence; and
c) analytical theme development to generate integrated interpretations that transcended individual study findings.
d) Cross-theme integration was achieved through conceptual mapping of pathways connecting cognitive, behavioral, and emotional constructs.

The methodological quality of the included studies was considered throughout the thematic synthesis. Findings from studies rated as high methodological quality were given greater interpretative weight when identifying consistent patterns and developing analytical themes, findings from moderate-quality studies were treated as suggestive of association rather than causation, and findings from low-quality studies were interpreted cautiously and used chiefly to identify candidate constructs. Non-empirical conceptual and theoretical contributions were retained for interpretive framing but were not treated as direct evidence of measured effects. This graded approach strengthened the credibility of the synthesized evidence across the cognitive, behavioral, and emotional domains.

## RESULTS

### Overview of Included Studies

The 57 included studies spanned a diverse methodological landscape and geographic distribution. Study designs included randomized controlled trials (n = 2), systematic reviews and meta-analyses (n = 11), cross-sectional surveys (n = 5), longitudinal surveys (n = 3), mixed-methods studies (n = 3), experimental laboratory studies (n = 7), conceptual and theoretical analyses (n = 22), and expert opinion syntheses (n = 4). Studies were conducted predominantly in the United States (n = 23), the United Kingdom (n = 4), India (n = 4), Germany (n = 4), China (n = 2), Australia (n = 2), the Netherlands (n = 2), and across multi-national or international samples (n = 17). Deployment contexts included higher education (n = 16), K–12 education (n = 4), healthcare and mental health (n = 15), organizational and professional settings (n = 6), consumer technology contexts (n = 10), and domestic or general daily use (n = 6). Sample sizes ranged from single-participant case studies to large-scale surveys involving up to 56,680 participants. AI modalities examined included intelligent tutoring systems, generative large language models, AI health chatbots, algorithmic social media platforms, social robots, voice assistants, and general AI advisory systems. The methodological quality of the 57 included studies was assessed using the Joanna Briggs Institute (JBI) Critical Appraisal Checklists appropriate to each study design. Of the included studies, 13 (23%) were rated as high quality, 19 (33%) as moderate quality, 4 (7%) as low quality, and 21 (37%) were conceptual, theoretical, or opinion-based sources for which formal empirical appraisal was not applicable.

Table 2: Methodological Quality Classification of the Included Studies

| **Theme** | **Total Sources** | **High Quality** | **Moderate Quality** | **Low Quality** | **Not Applicable ( Conceptual/ Theoretical)** |
|---|---|---|---|---|---|
| Theme A – Cognitive Impact | 18 | 4 | 4 | 1 | 9 |
| Theme B –Behavioural Impact | 19 | 4 | 5 | 3 | 7 |
| Theme C – Psychosocial Impact | 20 | 5 | 10 | 0 | 5 |
| TOTAL (all themes) | 57 | 13 | 19 | 4 | 21 |

Across the three themes, high-quality ratings were concentrated among large, well-controlled experimental studies and PRISMA-guided systematic reviews with meta-analysis, several of which reported dual independent screening, prospective registration, or convergent multi-instrument measurement. Moderate-quality studies were dominated by cross-sectional surveys and mixed-methods designs that used validated instruments but could not support causal inference. Low-quality ratings, concentrated in Theme B, primarily reflected high risk of bias in underlying primary studies, small or single-case samples, and low survey response rates; no low-quality sources were identified in Theme C. Recurring weaknesses across all three themes included a predominance of cross-sectional and self-report designs, limited or absent control conditions, under-reporting of blinding, reflexivity, and ethical approval, and, in Theme C specifically, undisclosed conflicts of interest among authors affiliated with the developers of the AI products being evaluated. These quality ratings informed the interpretation of evidence throughout the thematic synthesis.

**Theme A: Cognitive Impact of Artificial Intelligence Cognitive Offloading, Memory and Attention**

Across the 18 studies included in Theme A, cognitive offloading emerged as the most pervasive and consistently documented cognitive outcome, being addressed in 16 studies. Foundational experimental evidence demonstrated that the expectation of future online access reduced internal encoding of factual information by approximately 29% while strengthening memory for retrieval pathways, providing empirical support for the extended mind hypothesis (Sparrow et al., 2011). Subsequent evidence indicated that this mechanism extends to generative AI contexts. A PRISMA-compliant scoping review synthesizing 87 studies identified cognitive offloading as the dominant psychological theme in the generative AI literature, with dependence on AI enhancing short-term performance while diminishing autonomous cognitive processing (Ni & Li, 2026; Kasneci et al., 2023). Memory effects (9 studies) exhibited a characteristic dissociation, whereby AI assistance preserved or improved immediate task efficiency but compromised deeper encoding, semantic retention, and independent knowledge construction. Attentional effects (4 studies) were consistently adverse under prolonged AI use. Large-magnitude associations were reported between sustained AI interaction and attention strain ($r = .874$) as well as information overload ($r = .905$) (Shalu et al., 2025). These findings were reinforced by a systematic scoping review showing that AI use shortened students' attention spans and hindered sustained concentration on cognitively demanding tasks (Kundu & Bej, 2025; Zhai, 2023).

**Critical Thinking and Decision-Making**

Critical thinking (15 studies) and decision-making (13 studies) represented the most extensively investigated cognitive domains. The principal mechanism undermining critical thinking was automation bias, whereby humans' tendency to attribute meaning creates vulnerability to fluent yet semantically ungrounded AI outputs. The stochastic parrot's framework demonstrated that large language models manipulate linguistic form without genuine semantic understanding, a conclusion supported by empirical evidence of GPT-3's limitations in mathematical, semantic, and ethical reasoning (Bender et al., 2021; Floridi & Chiriatti, 2020). Repeated AI interaction was associated with lower self-assurance in independent decision-making ($r = -.360$) and reduced decision confidence ($r = -.401$), while a synthesis of 68 studies identified a broader process of cognitive disempowerment arising from the gradual relinquishment of autonomous reasoning (Shalu et al., 2025; Ani et al., 2025; Dwivedi et al., 2023). These effects were strongly moderated by contextual conditions. Explicit prompting for critical engagement with AI feedback preserved epistemic vigilance among graduate students (Olga et al., 2023). Conversely, a pedagogical paradox emerged whereby AI use without teacher mediation appeared more effective in fostering independent problem-solving than teacher-supported AI use, suggesting that excessive guidance may inadvertently encourage cognitive offloading rather than autonomous reasoning (Kundu & Bej, 2025).

**Creativity, Problem-Solving, and Metacognition**

Creativity findings (9 studies) were notably mixed and developmental-stage-dependent. Younger learners showed consistent creativity benefits through AI-assisted exploration (Kundu & Bej, 2025), while theoretical analyses raised substantive concerns about homogenization and derivative output in generative AI contexts

(Bender et al., 2021; Zhang & Magerko, 2025). Problem-solving evidence (10 studies) similarly indicated that AI supports macro-level problem structuring but creates risks of bypassing analytical reasoning when students use AI to generate rather than develop solutions (Tuomi, 2018). Metacognition (14 studies) emerged as simultaneously the most threatened and most protective cognitive capacity. The Eliza effect- systematic over-estimation of AI reliability and accountability (Zhang & Magerko, 2025)- represents the primary metacognitive failure mode introduced by generative AI. Luckin (2018) positioned metacognition as the most important element of human intelligence and one beyond AI's current replication, making its systematic cultivation an educational imperative. The evidence converged on AI literacy education- explicitly cultivating awareness of AI limitations, source evaluation, and skeptical engagement- as the primary protective factor (Tadimalla & Maher, 2024; Zhang & Magerko, 2025; Akgun & Greenhow, 2022; Shneiderman, 2022; Russell, 2019).

**Theme B: Behavioral Impact of Artificial Intelligence Technology Adoption and Trust**

Technology adoption (18 of 19 studies) functioned as the foundational behavioral gateway through which downstream effects of AI were mediated. Empirical evidence derived from structural equation modelling studies consistently revealed a dual-pathway architecture comprising a utilitarian route driven by perceived usefulness and performance expectancy, and a hedonic route influenced by perceived enjoyment and anthropomorphic characteristics (Chedrawi et al., 2025; Choudhary et al., 2024; Moussawi et al., 2020; Pillai & Sivathanu, 2020). Collectively, these findings suggest that AI adoption is shaped not only by functional value but also by users' affective responses and perceptions of social presence. Quantitatively, enabling factors were found to outweigh inhibitory influences by nearly sevenfold ($\beta = 0.657$ vs. $-0.096$), accounting for 58.8% of the variance in adoption intention and highlighting the predominance of facilitating mechanisms in determining behavioral acceptance of AI (Choudhary et al., 2024). Algorithmic trust (17 studies) occupied a pivotal moderating role and was characterized by a fundamental paradox. Under favorable conditions- objective tasks, no observed errors, transparent systems- people exhibit algorithm appreciation, weighting algorithmic advice more heavily than identical human advice (Weight on Advice: 0.45 vs. 0.30; Logg et al., 2019). Yet trust is fragile; Dietvorst et al.'s (2015) algorithm aversion paradigm demonstrated that a single observed failure reduced algorithmic choice from 70% to 35%, an asymmetric intolerance that does not apply to equivalent human error. Bonezzi et al. (2022) attributed this asymmetry to a projection bias- the illusion that one understands human decision-making better than algorithmic processes- a cognitive mechanism with important implications for appropriate reliance calibration.

**Learning Behavior, Engagement, and Productivity**

Learning behavior (11 studies) and engagement (11 studies) provided the most robustly supported positive behavioral outcomes identified across the review. The strongest empirical evidence for AI-enhanced learning behavior originated from decades of research on Intelligent Tutoring Systems (ITS), which consistently demonstrated improvements in learning performance and knowledge acquisition. Syntheses of more than 100 controlled ITS evaluations reported learning gains ranging from small to large effect sizes ($d = 0.3–1.2$) and reductions in time-to-mastery of 30–50%, representing the strongest and most consistent behavioral evidence base in the review (Woolf, 2010). These findings were further corroborated by systematic reviews of AI in education, which confirmed the effectiveness of adaptive and intelligent learning systems across diverse educational settings (Zawacki-Richter et al., 2019; Chassignol et al., 2018). Engagement trajectories in health AI contexts exhibited a non-monotonic pattern, with initial engagement rates of approximately 72% declining to 31% over eight months before rebounding at programmed termination, suggesting that sustained behavioral commitment requires active personalization mechanisms and continued user support (Aggarwal et al., 2023). Productivity gains (9 studies) were consistently documented across educational, health, and organizational contexts, reflecting AI's capacity to improve efficiency, reduce cognitive burden, and accelerate task completion. Adaptive learning systems and automated feedback mechanisms contributed to substantial reductions in learning time, while AI-assisted interventions in healthcare settings facilitated more efficient service delivery and behavioral support (Woolf, 2010; Aggarwal et al., 2023; Zawacki-Richter et al., 2019). Decision behavior (9 studies) emerged as a more complex and context-dependent construct. Experimental evidence demonstrated that preferences for algorithmic recommendations are strongly moderated by perceptions of task objectivity.

Objective domains promoted greater reliance on AI, whereas subjective domains involving personality, relationships, and identity elicited significantly greater resistance unless AI's affective capabilities were explicitly framed (Castelo et al., 2019). Complementing these findings, studies on algorithm appreciation and algorithm aversion showed that reliance on AI is shaped not only by objective performance but also by prior experiences and users' tolerance for algorithmic error, highlighting the importance of trust calibration in human–AI decision-making processes (Logg et al., 2019; Dietvorst et al., 2015).

**Human Agency and Dependency**

Human agency (14 studies) emerged as the most consistently at-risk behavioral construct. Mechanisms through which AI diminishes agency were varied: predictive algorithmic governance pre-empts students' behavioral change by assuming continuation of past patterns (Williamson & Eynon, 2020); emotional surveillance systems incentivize performative compliance; expert over-reliance reduces vigilant information-seeking; and educational systems designed without pedagogical co-creation systematically displace teacher professional judgment (Holmes et al., 2022). Dependency (3 studies), while underrepresented empirically, was identified across multiple review studies as an emergent endpoint of sustained AI engagement, particularly where AI replaces rather than augments cognitive processes.

**Motivation, Decision Behavior, and Behavioral Change**

Motivation (10 studies), decision behavior (9 studies), and behavioral change (3 studies) together represent the more proximal outcomes of AI-mediated behavioral influence. Motivation emerged as a consistently positive construct across educational, health, and consumer contexts, driven primarily by AI's capacity to provide personalized feedback, adaptive support, and reduced uncertainty. Evidence from intelligent tutoring systems demonstrated reductions in mathematics anxiety and improvements in learner confidence, while adaptive learning systems promoted persistence and self-regulation (Woolf, 2010; Luckin et al., 2016). In health contexts, chatbots employing motivational interviewing techniques strengthened health-related motivation, and hedonic motivation emerged as one of the strongest enablers of AI adoption intentions (Aggarwal et al., 2023; Choudhary et al., 2024). Reviews of generative AI applications further indicated that personalized responses and immediate feedback increased learner motivation and participation (Montenegro-Rueda et al., 2023; Fraiwan & Khasawneh, 2023). Collectively, these findings suggest that AI-supported motivation operates through competence enhancement and autonomy support, thereby encouraging sustained engagement with AI systems. Decision behavior was characterized by the competing phenomena of algorithm appreciation and algorithm aversion. Experimental evidence demonstrated that individuals weighted algorithmic advice more heavily than identical human advice, with 88% preferring algorithmic recommendations over human judgment (Logg et al., 2019). However, this preference proved highly context dependent. Objective tasks promoted algorithm reliance, whereas subjective domains involving personality, relationships, and identity elicited greater resistance (Castelo et al., 2019). Furthermore, exposure to a single algorithmic error reduced algorithm choice from approximately 70% to 35%, irrespective of the algorithm's superior performance, illustrating the phenomenon of algorithm aversion (Dietvorst et al., 2015). Additional evidence from educational and organizational contexts suggested that automation complacency, predictive labelling, and reduced vigilance may further shape AI-supported decision-making (Dwivedi et al., 2021; Baker & Inventado, 2014; Williamson & Eynon, 2020). Together, these findings indicate that decision behavior is governed not only by algorithmic accuracy but also by trust calibration, task characteristics, and prior experiences with AI systems. Behavioral change, although represented in only three studies, constituted the most outcome-oriented dimension of behavioral impact. Evidence derived primarily from health AI applications, where improvements were observed in physical activity, dietary practices, smoking cessation, medication adherence, and help-seeking behaviors (Aggarwal et al., 2023). These effects were attributed to the non-judgmental and continuously available nature of AI chatbots, which reduced barriers to disclosure and promoted sustained participation. Responsible AI practices were also shown to positively influence users' behavioral intentions, while generative AI tools appeared capable of facilitating changes in study habits and learning practices (Chedrawi et al., 2025; Montenegro-Rueda et al., 2023; Luckin et al., 2016). Although the current evidence base remains limited and largely observational, the findings indicate that AI has

considerable potential to drive meaningful behavioral transformation, particularly when interventions are personalized and sustained over time.

**Theme C: Emotional Impact of Artificial Intelligence Anxiety, Stress, and Mental Health**

Mental health (15 studies), anxiety (10 studies), and stress (6 studies) were the most extensively investigated emotional constructs, revealing markedly divergent patterns across therapeutic and social AI environments. The strongest and most consistent evidence concerned the effects of clinical chatbots on depressive symptoms. Findings from three randomized controlled trials and two systematic reviews with meta-analyses demonstrated that structured AI-delivered interventions based on cognitive behavioral therapy, dialectical behavior therapy, mindfulness, and integrative approaches produced significant reductions in depressive symptoms compared with passive controls, with effect sizes ranging from $d = 0.44$ to $d = 0.68$ in individual studies and a pooled effect size of SMD = −0.55 (95% CI [−0.87, −0.23], $p < .001$) across studies (Abd-Alrazaq et al., 2020; Du et al., 2025; Fitzpatrick et al., 2017; Fulmer et al., 2018; Inkster et al., 2018). An updated meta-analysis further distinguished between rule-based and large language model (LLM)-based chatbots, reporting significant effects for rule-based systems ($g = 0.266$, $p = .04$) but insufficient evidence for LLM-based interventions, challenging assumptions that greater technological sophistication necessarily translates into superior clinical outcomes (Du et al., 2025; Miner et al., 2019; Laranjo et al., 2018). In contrast, evidence from AI-driven social media environments revealed a markedly different emotional profile. Problematic engagement with algorithmically curated platforms was associated with higher levels of depression, anxiety, and social isolation, with effects mediated primarily through the displacement of real-world social support (Meshi & Ellithorpe, 2021). These outcomes appear to be rooted in platform architectures designed to maximize engagement, including personalized feeds, push notifications, and endless scrolling, which inadvertently foster anxiety, emotional dependency, and social comparison processes (Montag & Hegelich, 2020).

**Emotional Support, Self-Efficacy, and Well-Being**

Fourteen studies examined emotional support, collectively demonstrating that AI systems can provide emotionally meaningful experiences even when users are fully aware that they are interacting with non-human agents. This phenomenon is consistent with the media equation framework, which posits that humans instinctively apply social rules to sufficiently social technologies (Nass & Reeves, 1996, as operationalized in Broadbent, 2017). Evidence synthesized across mental health studies further indicated that patients with major depressive disorder reported stronger therapeutic alliances with conversational AI than with human clinicians in some contexts (Vaidyam et al., 2019). Likewise, a large multinational study involving 270 participants from 29 countries found that 99.6% of ChatGPT users seeking emotional support rated the experience as at least somewhat helpful, with emotional safety, validation, connectedness, and relief emerging as dominant themes (Luo et al., 2025). Three mechanisms appeared to underlie the effectiveness of AI-mediated emotional support: reduced fear of social judgment, continuous availability through 24-hour access and routine check-ins, and therapeutic process factors such as empathy, accountability, and skill transmission that resemble the common factors underlying human psychotherapy (Vaidyam et al., 2019; Luo et al., 2025; Calvo & Peters, 2014). Improvements in self-efficacy (7 studies) were consistently reported in clinical contexts, with users describing enhanced coping strategies, emotional insight, and self-awareness. Well-being outcomes (8 studies), however, were more heterogeneous. While therapeutic chatbots produced modest improvements, AI use in social media contexts was associated with poorer well-being, suggesting that emotional benefits are fundamentally design dependent. Ethical shortcomings, including algorithmic bias, privacy concerns, and lack of transparency, may further undermine trust and diminish AI's capacity to promote psychological well-being (Lee et al., 2021; Calvo & Peters, 2014).

**Social Comparison, Loneliness, and Human Relationships**

The most consequential and temporally complex emotional findings concerned loneliness (8 studies) and human relationships (10 studies). Social robots demonstrated clinically meaningful reductions in loneliness, particularly among older adults in institutional care settings, with randomized controlled studies reporting decreased social

isolation and agitation among patients with dementia (Broadbent, 2017). However, the strongest methodological evidence in the review- a 12-month longitudinal study involving 2,149 participants across four waves- revealed a contrasting pattern, with social chatbot use predicting increased emotional isolation over time ($\beta = 0.07–0.08$, $p = .006$), a finding that remained robust across all six sensitivity analyses (Folk & Dunn, 2026). This temporal reversal, whereby short-term emotional comfort may contribute to longer-term loneliness, represents one of the most important findings within Theme C. Clinician perspectives further reinforced these concerns. General practitioners in the United Kingdom overwhelmingly viewed empathy, therapeutic communication, and patient-centered care as fundamentally human capacities that AI cannot adequately reproduce, emphasizing that medicine remains "a people business" in which interpersonal relationships are central (Blease et al., 2019). Complementing these concerns, ethnographic evidence suggested that sustained interaction with AI may simultaneously lower expectations of human relationships while increasing emotional investment in machines, creating a dynamic in which low-cost and low-risk emotional support reduces motivation to maintain more demanding but ultimately rewarding human connections (Turkle, 2011). This mechanism is consistent with reinforcement architectures embedded within digital platforms and points to a potentially addictive dimension of AI-mediated social interaction (Montag & Hegelich, 2020; Kretzschmar et al., 2019).

**Emotional Dependency and Emotional Adaptation**

Emotional dependency (9 studies) emerged as a clinically significant yet systematically underreported consequence of prolonged AI interaction. Longitudinal evidence from domestic social robot users demonstrated the development of genuine emotional attachments over six months, with participants often displaying reluctance to acknowledge these relationships, suggesting that self-report measures may underestimate the prevalence of dependency (De Graaf et al., 2016). Extending these findings, longitudinal evidence indicated that AI companionship may function as a compensatory strategy that ultimately becomes self-defeating. Users appeared to adapt to the predictability, availability, and non-judgmental support provided by AI companions, gradually perceiving the uncertainty and emotional demands inherent in human relationships as less attractive despite their greater long-term benefits (Folk & Dunn, 2026). In contrast, emotional adaptation represented the positive counterpart to dependency and was most evident in therapeutic contexts, where structured AI interactions facilitated the acquisition of adaptive coping skills and emotional self-regulation. The findings collectively suggest that the emotional consequences of AI depend critically on system design. Capacity-building systems, characterized by skill transmission, autonomy support, and explicitly time-limited interactions, appear to promote emotional resilience, whereas demand-satisfying systems that provide unlimited companionship and continuous validation may foster maladaptive attachment patterns. This distinction highlights the importance of ethical design principles and carries significant implications for the development of emotionally intelligent AI systems (Fiske et al., 2019).

**Cross-Theme Integration: The AI-Psychology Cascade**

The findings across the three thematic domains are not independent but constitute an integrated psychological cascade in which cognitive adaptations produce behavioral modifications that generate emotional consequences, with feedback loops operating in both directions:

AI Exposure → Cognitive Adaptation → Behavioral Modification → Emotional Consequences → Long-term Human Functioning

At the cognitive level, AI exposure reshapes attentional processes, memory strategies, and metacognitive regulation. The cognitive habit of offloading, although rational in the short term, fosters a behavioral pattern of deference that undermines the motivational and developmental foundations of autonomous cognition. Behavioral dependency, once established, produces emotional consequences: relief from the anxiety of autonomous decision-making, which reinforces further reliance; decreased tolerance for cognitive effort, which shapes future learning behavior; and reduced investment in demanding human relationships, which amplifies loneliness. The emotional state of anxiety, in turn, accelerates AI-seeking as a compensatory strategy, deepening the cycle. This cascade is not unidirectional or deterministic. At each level, moderating factors- AI literacy,

pedagogical design, system transparency, user metacognitive capacity; can redirect the cascade toward enhancement rather than dependency. The evidence suggests that the cascade is most likely to be adaptive when AI is explicitly framed as a scaffolding tool requiring active critical engagement, and most likely to be maladaptive when AI is used passively, uncritically, and in substitution for human cognitive and relational effort.

**Table 3: Integrated Cross-Theme Findings**

| Cross-Theme Pathway | Cognitive Mechanism | Behavioural Outcome | Emotional Consequence |
|---|---|---|---|
| **AI Exposure → Cognitive Offloading** | Reduced internal encoding; transactive memory shift to AI | Increased adoption; reduced critical engagement | Relief from cognitive strain; risk of anxiety when AI unavailable |
| **Automation Bias → Trust Miscalibration** | Epistemic passivity; over-attribution of AI reliability | Algorithm appreciation then aversion; decision quality variance | Frustration, distrust, or false reassurance |
| **Agency Erosion → Dependency** | Metacognitive atrophy; loss of self-regulatory capacity | Reduced autonomous learning; performative compliance | Anxiety, lowered self-efficacy, emotional dependency |
| **AI Therapeutic Engagement** | Structured CBT delivery; skill acquisition through AI interaction | Behavioural activation; adaptive help-seeking | Reduced depression/stress; improved self-efficacy and well-being |
| **Social Media AI → Comparison** | Attentional fragmentation; shallow information processing | Compulsive checking; displacement of human interaction | Upward social comparison; anxiety; loneliness; reduced WB |
| **AI Literacy Mediation** | Metacognitive monitoring; epistemic vigilance | Critical engagement with AI; reduced over-reliance | Preserved emotional regulation; adaptive AI use patterns |
| **Companionship AI → Relational Displacement** | Reduced elaborative processing of human social cues | Reduced investment in human relationship maintenance | Increased loneliness over 12 months; emotional impoverishment |
| **ITS Learning Gains** | Personalised scaffolding; error modelling; adaptive pacing | Improved learning behaviour; productivity gains d=0.3–1.2 | Increased motivation; reduced anxiety; enhanced self-efficacy |

# DISCUSSION

## Interpretation of Major Findings

The present systematic review synthesized evidence from 57 peer-reviewed studies to establish that artificial intelligence exercises substantial, multidimensional, and contextually contingent effects on human cognition, behavior, and emotional well-being. The overarching finding is not that AI is uniformly beneficial or harmful, as the evidence supporting both trajectories is considerable, but rather that the direction, magnitude, and durability of AI's psychological effects are substantially malleable through design choices, educational interventions, and policy frameworks. This conclusion carries profound implications for how AI systems are developed, deployed, evaluated, and regulated. The interpretation of the findings should be considered in light of the methodological quality of the included studies. Although approximately one-quarter of the evidence base comprised studies rated as high methodological quality and a further one-third as moderate quality, the predominance of cross-sectional designs and self-reported outcome measures limited the ability to draw causal inferences. In addition, a proportion of studies within the emotional and psychosocial domain (Theme C)

involved industry-affiliated efficacy evaluations, which may introduce potential conflicts of interest and should therefore be interpreted with appropriate caution. Consequently, the evidence provides moderate confidence in the observed associations between AI use and psychological outcomes, whereas causal claims remain tentative and require confirmation through well-designed longitudinal, experimental, and mixed-methods research.

Three major findings require particular interpretive attention. First, cognitive offloading is the most robust and pervasive cognitive mechanism through which AI reshapes human psychology. The evidence from Sparrow et al. (2011) through to Ni and Li (2026) establishes a clear mechanistic chain: AI availability attenuates the motivational incentive for internal encoding → attentional resources are reoriented toward external retrieval pathways → metacognitive vigilance is reduced → autonomous cognitive capacity atrophies. This mechanism, originally demonstrated for factual recall, has now been shown to generalize to the full complexity of generative AI tasks including writing, problem-structuring, and decision support. The civilizational risk articulated by Russell (2019) and Tuomi (2018) that large-scale cognitive offloading may progressively erode the distinctively human cognitive capacities that enable autonomous flourishing, is not merely speculative but is grounded in empirically documented mechanisms. Second, the paradox of trust and agency constitutes the central behavioral tension in the evidence base. Human beings are simultaneously prone to algorithm appreciation (Logg et al., 2019), particularly for objective, low-stakes decisions, and algorithm aversion (Dietvorst et al., 2015), triggered by observed failure or subjective task framing. Neither extreme represents appropriate trust calibration, and both produce suboptimal behavioral outcomes: over-trust fosters automation complacency and agency erosion; under-trust leads to the systematic rejection of superior AI recommendations. The cognitive mechanism underlying this paradox, the projection bias identified by Bonezzi et al. (2022) suggests that lay beliefs about algorithmic versus human cognition are systematically mis-calibrated in ways that cannot be corrected by mere familiarity with AI systems. Explicit trust calibration education may be necessary. Third, the temporal inversion documented in the emotional domain, where AI interventions that provide immediate comfort predict increased loneliness over 12 months represents a finding of major practical importance (Folk & Dunn, 2026). The psychological mechanism proposed is habituation to the predictable, non-demanding, and non-vulnerable character of AI social interaction, which progressively reduces motivation to invest in the more difficult but more rewarding forms of human connection. This mechanism is structurally analogous to other forms of hedonic adaptation and may be equally resistant to volitional interruption.

**Comparison with Existing Theories**

The findings of this review may be interpreted through multiple established theoretical frameworks, each of which captures a different dimension of AI's psychological impact. Cognitive Load Theory (Sweller, 1988) provides partial but limited explanatory purchase. The observation that AI can reduce extraneous cognitive load by minimizing irrelevant processing and tailoring instruction to learners' needs suggests that cognitive resources may be freed for meaningful learning. This mechanism is consistent with the positive educational outcomes documented in intelligent tutoring system research (Woolf, 2010) and more recent evidence on adaptive AI-based learning environments (Kundu & Bej, 2025). However, the theory does not account for the motivational and metacognitive consequences of persistent offloading, nor does it address the atrophy of intrinsic cognitive processing that appears to result from sustained AI use. Clark and Chalmers' (1998) Extended Mind Theory, as applied by Sparrow et al. (2011), provides a mechanistic foundation for cognitive offloading. Their findings show that humans treat digital systems as transactive memory partners, a strategy that is cognitively rational in the short term but may carry long-term costs as reliance on external systems expands with AI capability. The critical extension required by the present evidence base is temporal: the theory predicts stable integration of external cognitive resources but does not account for the developmental consequences of never building the internal cognitive structures that AI now provides. Social Cognitive Theory (Bandura, 1986) offers the most coherent framework for understanding AI's behavioral and emotional effects, particularly through the constructs of self-efficacy and observational learning. The self-efficacy benefits documented in therapeutic AI trials (Fitzpatrick et al., 2017; Fulmer et al., 2018) are consistent with Bandura's model of efficacy-building through mastery experience and social persuasion. The erosion of human agency documented in educational and organizational AI contexts is, conversely, a case study in the destruction of self-efficacy through the progressive replacement of mastery experience with algorithmic substitution. Self-Determination Theory (Deci & Ryan,

1985) illuminates both the motivational benefits and the dependency risks of AI. Motivational enhancement through AI occurs when systems satisfy basic psychological needs for competence (through personalized feedback), autonomy (through adaptive choice), and relatedness (through social AI). Motivational atrophy and dependency emerge when AI satisfies these needs in ways that do not build the intrinsic motivational capacities that sustain autonomous functioning, a distinction precisely captured by the contrast between capacity-building and demand-satisfying AI identified in Theme C. The Technology Acceptance Model (Davis, 1989) and its extensions (UTAUT) explain technology adoption pathways (Chedrawi et al., 2025; Choudhary et al., 2024; Pillai & Sivathanu, 2020) but do not address the downstream psychological consequences of adoption- a limitation that underscores the need for post-adoption longitudinal research. Media Dependency Theory (Ball-Rokeach & DeFleur, 1976), while originally developed for mass media, provides a compelling framework for understanding the functional displacement of human social connection by AI: as users become dependent on AI-mediated social information and emotional support, the structural conditions for relational atrophy and loneliness are established.

**Positive Versus Negative Psychological Outcomes**

A schematic summary of AI's net psychological direction across domains reveals a pattern that defies simple characterization as either beneficial or harmful. Positive outcomes predominate in structured, time-limited, skill-building contexts: ITS-mediated learning gains (d = 0.3-1.2), productivity enhancement, motivation improvement through personalized feedback, therapeutic depression reduction through CBT-based chatbots, and stigma reduction facilitating access to care. Negative outcomes predominate in unstructured, prolonged, passive-use, or substitutive contexts: cognitive offloading leading to attentional fragmentation and metacognitive atrophy, algorithmic trust miscalibration, human agency erosion, social media-driven anxiety and social comparison, and companionship AI-driven loneliness amplification. This pattern suggests that the central design question for AI systems is not whether to integrate AI into human environments but under what conditions AI augments versus substitutes human capacity, and whether the conditions of deployment are those that produce augmentation or substitution. The evidence indicates that augmentation is substantially more likely when AI is transparent and explainable, explicitly time-limited, structured around skill transmission rather than demand satisfaction, accompanied by metacognitive literacy instruction, and embedded within human oversight frameworks.

**Contextual Differences**

The direction and magnitude of AI's psychological effects were substantially moderated by several contextual factors across the three thematic domains. Age and developmental stage exerted a consistent moderating influence. Younger learners (preschool to primary) showed greater cognitive and creative benefits from AI interaction, attributable to reduced anxiety, greater novelty response, and lower prior cognitive investment in the functions AI supports (Kundu & Bej, 2025). Older learners and adult populations showed greater risks of cognitive offloading, critical thinking erosion, and metacognitive atrophy, possibly because they have more established cognitive habits that AI interaction disrupts. In emotional domains, older adults showed the most robust loneliness reduction from social robots (Broadbent, 2017), while young adults showed the greatest vulnerability to social media AI-driven anxiety and depression (Meshi & Ellithorpe, 2021). AI literacy and metacognitive preparation were the most consistently identified protective factors across all three domains. Users with higher AI literacy demonstrated more critical engagement with AI outputs, more appropriate trust calibration, and more adaptive emotional responses to AI interaction. This finding converges across cognitive (Zhang & Magerko, 2025; Tadimalla & Maher, 2024), behavioral (Holmes et al., 2022), and emotional (Kretzschmar et al., 2019) domains, suggesting that AI literacy is not a domain-specific competency but a cross-cutting psychological protective factor. Exposure duration and intensity demonstrated particularly complex moderating effects. Short-term AI exposure consistently produced positive outcomes across all domains: efficiency gains, learning improvements, emotional relief. Long-term, intensive exposure, particularly to unstructured or companionship-oriented AI, showed the opposite pattern: cognitive atrophy (Shalu et al., 2025; Ani et al., 2025), engagement decay (Aggarwal et al., 2023), and loneliness amplification (Folk & Dunn, 2026). This temporal reversal constitutes one of the most important and clinically relevant findings of the review.

## Emerging Integrated Psychological Model

Based on the cross-domain evidence synthesis, this review proposes an Integrated Psychological Model of AI Impact in which AI functions across four context-dependent roles that are not mutually exclusive but are substantially determined by design and deployment conditions:

As a Cognitive Enhancer, AI functions in structured, literacy-mediated, pedagogically designed contexts to offload routine processing, provide immediate corrective feedback, scaffold complex tasks, and free attentional resources for higher-order reasoning and creative elaboration. This role is most strongly evidenced in younger learner populations and explicitly designed human-AI collaborative systems. As a Behavioral Modifier, AI reshapes engagement patterns, learning habits, motivational architectures, and decision strategies through personalized feedback and adaptive incentive structures. This role produces net positive behavioral outcomes when AI systems are transparent, ethically governed, and supplementary to rather than substitutive of human judgment. As an Emotional Regulator, AI provides accessible, stigma-reduced, consistently available emotional support that produces genuine psychological benefits- particularly for depression, stress, and self-efficacy- in structured clinical contexts. This role is most effectively realized through time-limited, capacity-building, skill-transmitting AI systems that explicitly direct users toward human connection and professional care. As a Dependency Mechanism, AI creates structural conditions of cognitive, behavioral, and emotional dependency under conditions of prolonged, unregulated, passive, and substitutive use. The dependencies formed in each domain are mutually reinforcing: cognitive offloading attenuates metacognitive vigilance, which enables uncritical behavioral adoption, which generates emotional over-reliance, which further reduces motivation for the autonomous cognitive effort that could interrupt the cycle. This represents the primary systemic risk of AI integration at scale. The evidence suggests that which of these four roles predominates is not technologically determined but is substantially malleable through human choices: the design decisions of AI developers, the pedagogical frameworks of educators, the ethical governance of policymakers, and the metacognitive strategies of individual users. The paramount implication is that AI's psychological impact is neither inevitable nor unpredictable- it is designed.

## Implications

### Educational Implications

The evidence reviewed here carries urgent implications for educational practice at every level. Educators must recognize that AI tools are not pedagogically neutral; their integration into learning environments reshapes cognitive processes, behavioral habits, and emotional states in ways that demand deliberate, evidence-informed pedagogical design. Three educational priorities emerge with particular urgency. First, AI literacy should be developed as a cross-curricular competency rather than being confined to a specialist technical domain. The metacognitive capacities essential for effective and responsible AI use- including the critical evaluation of AI-generated outputs, recognition of algorithmic limitations and biases, calibrated trust in AI systems, and epistemic vigilance- require explicit and systematic instruction rather than being assumed to emerge through mere exposure to AI technologies. Consequently, AI literacy should be intentionally integrated across subject areas and educational stages. Existing frameworks for AI literacy provide practical guidance for curriculum design and implementation and can be adapted to suit different grade levels and disciplinary contexts (Zhang & Magerko, 2025; Tadimalla & Maher, 2024). Second, assessment practices must evolve to evaluate the higher-order cognitive processes that AI cannot readily substitute, including critical analysis, creative transformation, metacognitive reflection, and ethical reasoning, rather than focusing primarily on information products that AI systems can easily generate. This shift requires the redesign of assessment frameworks to prioritize authentic, process-oriented, and performance-based tasks that capture learners' reasoning, judgment, and reflective capacities. Contemporary scholarship offers practical guidance for reimagining assessment in the age of AI, emphasizing the need for systematic and proactive institutional reforms rather than reactive adaptations to technological change (Zhai, 2023; Kasneci et al., 2023). Third, teacher professional development must prepare educators to integrate AI tools in ways that enhance, rather than replace, students' cognitive development. Effective AI integration requires teachers to understand how AI can be used as a pedagogical support that fosters

critical thinking, problem-solving, creativity, and self-regulated learning, rather than encouraging passive dependence on automated outputs. Evidence suggests that even teacher-scaffolded AI use may inadvertently promote cognitive offloading if instructional strategies are not carefully designed, highlighting the importance of equipping educators with the knowledge and skills to distinguish between AI applications that augment student cognition and those that substitute for essential learning processes (Kundu & Bej, 2025).

### Policy Implications

The evidence reviewed here supports several specific policy directions. AI-driven platform features that exploit attentional and emotional mechanisms such as infinite scrolling, push notifications, algorithmically curated social feeds. should be subject to regulatory scrutiny proportionate to their documented psychological harms, particularly for young users. The EU AI Act's risk classification framework provides a regulatory model that could be extended to include psychological harm as a criterion for higher-risk AI classification. Algorithmic transparency requirements should be strengthened: the consistently documented link between AI opacity and both agency erosion and trust miscalibration indicates that explainability is not merely a technical preference but a psychological necessity. Policy frameworks that mandate explainability for high-stakes AI systems should be extended and enforced.

### Mental Health Implications

The mental health implications of this review are twofold. First, evidence indicates that structured AI-delivered therapeutic interventions can effectively reduce symptoms of depression among individuals with subclinical mental health concerns while simultaneously addressing barriers to care through reduced stigma, improved accessibility, and continuous availability of support (Laranjo et al., 2018; Miner et al., 2019). These advantages present a significant opportunity to extend the reach of mental health services, particularly for underserved and hard-to-reach populations. However, this potential must be pursued within clearly defined ethical boundaries. AI-delivered mental health support should be explicitly positioned as a complement to, rather than a replacement for, professional human care (Fiske et al., 2019). Furthermore, crisis detection and response protocols must be rigorously developed, validated, and regularly tested to ensure user safety (Laranjo et al., 2018; Miner et al., 2019). In addition, the risks of emotional dependency and overreliance on AI companions, as highlighted by recent research, should be proactively mitigated through system designs that strengthen human relationships and encourage timely engagement with professional and social support networks rather than substituting for them (Folk & Dunn, 2026; Fiske et al., 2019). The mental health risks of AI-driven social media platforms are sufficiently established to warrant clinical attention. Mental health professionals should routinely assess clients' social media use patterns, with particular attention to algorithmic engagement features that may be maintaining or amplifying anxiety, depression, and social isolation.

### Responsible AI Use

The evidence reviewed here supports the principles of Human-Centered AI (Shneiderman, 2022) as an ethical and empirical framework for AI development: AI systems that preserve human agency, provide transparent explanations, enable user oversight, and explicitly support rather than replace human cognitive and relational capacities are not merely ethically preferable but psychologically superior to their autonomous alternatives. The positive computing framework proposed by Calvo and Peters (2014) designing AI to cultivate emotional intelligence, self-insight, and prosocial motivation should become a standard design criterion alongside performance and efficiency.

## RECOMMENDATIONS FOR PRACTICE

1. Implement AI literacy curricula across educational levels, explicitly cultivating metacognitive evaluation skills
2. Redesign assessments to evaluate human-specific cognitive capacities that AI cannot substitute
3. Develop and enforce algorithmic transparency standards for high-stakes AI systems

4. Establish evidence-based guidelines for AI chatbot deployment in mental health contexts
5. Incorporate AI use pattern assessment into routine mental health clinical practice
6. Design AI systems with explicit mechanisms for building human connection rather than substituting for it
7. Mandate longitudinal impact assessments for AI systems deployed at scale in educational and health contexts

## LIMITATIONS

Several limitations of this review should be acknowledged. Although methodological quality was systematically appraised across all three themes using the JBI Critical Appraisal Checklists, the evidence base was dominated by cross-sectional and self-report designs, and 21 of the 57 included sources (37%) were conceptual, theoretical, or opinion-based contributions without primary empirical data. In addition, several of the highest-profile efficacy trials in Theme C disclosed conflicts of interest linked to the developers of the technologies evaluated, and variability in study designs, AI applications, and outcome measures introduced considerable heterogeneity across the included studies, limiting the strength of any pooled or comparative conclusions. Some other notable limitations are:

First, the inclusion of both empirical studies and conceptual analyses within a single synthesis introduces the possibility of conflating different levels of evidence. While theoretical contributions provide valuable conceptual frameworks for interpreting the emerging psychological implications of AI (Russell, 2019; Tuomi, 2018; Turkle, 2011), they do not possess the same evidentiary strength as empirical investigations. This methodological heterogeneity complicates broad claims regarding the overall strength of evidence. To address this concern, the review explicitly classified studies according to methodological quality and consistently differentiated theoretical propositions from empirically supported findings throughout the Results and Discussion. Second, the rapid pace of AI development limits the temporal generalizability of the reviewed evidence. Earlier investigations examining technologies such as GPT-2 and traditional Intelligent Tutoring Systems (ITS) were conducted within technological contexts that differ substantially from contemporary large language model (LLM)-based AI systems. As AI capabilities evolve, so too do their potential cognitive, educational, and psychological impacts. Recent evidence indicates that LLM-based mental health chatbots continue to lack robust evidence of therapeutic efficacy, demonstrating that successive generations of AI require independent evaluation rather than extrapolation from earlier technologies (Du et al., 2025). Accordingly, the conclusions of this review should be interpreted as applying primarily to AI systems deployed through 2025. Third, the existing evidence base remains heavily concentrated within Western, Educated, Industrialized, Rich, and Democratic (WEIRD) populations, thereby limiting the external validity of current findings. The limited attention to equity and inclusion in AI psychology research constrains understanding of AI's psychological effects across diverse socioeconomic groups, linguistic communities, individuals with disabilities, and geographically underserved populations (Ni & Li, 2026). Consequently, the generalizability of existing findings beyond WEIRD contexts remains limited and represents an important priority for future research. Fourth, longitudinal evidence remains considerably less developed than the predominantly cross-sectional literature. Many of the most significant psychological consequences of AI- including cognitive dependency, cognitive atrophy, and relational displacement- are developmental processes that unfold gradually over time and therefore require longitudinal designs for rigorous examination. Although emerging evidence has begun documenting these longer-term concerns, the relative scarcity of longitudinal research limits current understanding of the durability, progression, and cumulative effects of AI-mediated psychological change (Folk & Dunn, 2026; Shalu et al., 2025). Finally, publication bias may have resulted in an overestimation of the positive effects of AI-based interventions, particularly within the literature on mental health chatbots. Documented developer conflicts of interest and the potential for selective reporting raise concerns regarding the objectivity of reported outcomes, suggesting that claims of therapeutic effectiveness should be interpreted with appropriate caution until supported by a broader body of independent, methodologically rigorous evaluations (Abd-Alrazaq et al., 2020).

### Future Research Directions

The evidence base reviewed here identifies several high-priority directions for future research that would substantially advance the field's capacity to understand and mitigate AI's psychological impact. Longitudinal studies with multi-wave designs tracking individuals from initial AI adoption through sustained use to potential dependency are urgently needed across all three thematic domains. The field's current reliance on cross-sectional and short-term experimental evidence has systematically obscured the temporal dynamics that are most consequential for human psychological well-being. Minimum follow-up periods of 12–24 months are recommended, consistent with Folk and Dunn's (2026) design, to capture the temporal inversion of AI's emotional effects. Experimental designs that manipulate specific design features of AI systems- transparency, time-limitation, skill-building orientation, versus companionship orientation- and measure their psychological consequences would allow causal attribution of the design-dependent effects documented in this review. Such designs would directly test the central claim of the integrated model: that AI's psychological role is designed rather than determined. Cross-cultural and equity-focused research is urgently needed to examine how socioeconomic status, language background, educational access, disability, and cultural context moderate psychological responses to AI. The current WEIRD bias in the evidence base prevents confident generalization to the majority of the world's AI users. School and workplace contexts each require dedicated research programs that account for the specific institutional dynamics, power relations, and developmental stakes that shape AI's psychological impact in these environments. The documented risks of algorithmic governance in educational institutions (Williamson & Eynon, 2020) and workplace automation complacency (Dwivedi et al., 2021) deserve rigorous primary empirical investigation rather than reliance on review-level evidence. Research on the ethical dimensions of AI emotional systems- particularly the conditions under which therapeutic AI creates pathological dependency, and the design interventions that can build emotional capacity rather than substituting for it- represents one of the most pressing practical research needs identified by this review. Fiske et al.'s (2019) conceptual framework provides a starting point for empirically grounded ethical research in this domain.

## CONCLUSION

This systematic review synthesized evidence from 57 peer-reviewed studies to address four research questions concerning AI's impact on human cognition, behavior, emotional well-being, and the integrated psychological mechanisms through which these effects propagate. The answers to these questions paint a complex but interpretable picture.

In response to RQ1, AI influences cognitive functioning primarily through cognitive offloading- a mechanism that simultaneously preserves short-term efficiency and erodes the long-term motivational and developmental foundations of autonomous cognitive capacity. The effects on critical thinking, metacognition, and decision-making autonomy are contingent upon the presence or absence of AI literacy education and critical engagement practices. In response to RQ2, AI influences behavioral patterns through a dual-pathway adoption architecture driven by utilitarian and hedonic motivation, producing robust learning, productivity, and motivation gains in structured contexts, while simultaneously fostering algorithmic trust miscalibration, agency erosion, and emergent dependency in unregulated or substitutive deployment conditions. In response to RQ3, AI affects emotional well-being through contextually divergent pathways: structured therapeutic AI reduces depression and stress while improving self-efficacy; AI-driven social media platforms amplify anxiety, social comparison, and loneliness; and companionship-oriented AI provides immediate emotional relief that predicts increased emotional isolation over 12 months, a temporal inversion of particular clinical importance. In response to RQ4, the integrated analysis reveals an AI-psychology cascade in which cognitive adaptations generate behavioral modifications that produce emotional consequences, with bidirectional feedback loops operating across all three domains. This cascade may be adaptive- when AI functions as a cognitive enhancer, behavioral facilitator, and emotional regulator under conditions of literacy, transparency, and human-centered design- or maladaptive, when AI functions as a dependency mechanism under conditions of uncritical, prolonged, and substitutive use. The defining implication of this review is that AI's psychological impact is neither inevitable nor unpredictable: it is designed. The cognitive, behavioral, and emotional consequences documented here are not inherent properties of AI technology but rather arise from the interaction between AI design, deployment context, user

literacy, and structural access conditions. This conclusion places the responsibility for AI's psychological effects squarely within the domain of human choice- the choices of developers, educators, policymakers, and users- and thereby within the domain of human agency. Ensuring that AI strengthens rather than supplants human psychological capacity is the defining challenge and the defining opportunity of the AI age.

## REFERENCES

1. Abd-Alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P., & Househ, M. (2020). An overview of the features of chatbots in mental health: A scoping review. International Journal of Medical Informatics, 132, 103978. https://doi.org/10.1016/j.ijmedinf.2019.103978
2. Aggarwal, A., Tam, C. C., Wu, D., Li, X., & Qiao, S. (2023). Artificial intelligence-based chatbots for promoting health behavioral changes: Systematic review. Journal of Medical Internet Research, 25, e40789. https://doi.org/10.2196/40789
3. Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing societal and ethical challenges in K-12 settings. AI and Ethics, 2(3), 431 440. https://doi.org/10.1007/s43681-021-00096-7
4. Ani, F., Hamzah, S., Damin, Z. A., Rameli, N., & Farid, S. (2025). Understanding the Human-AI Interaction: Psychological and Social Impacts of AI Integration in Daily Life. Journal of Techno-Social, 17(2), 126-133. https://publisher.uthm.edu.my/ojs/index.php/JTS/article/view/23882
5. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. *JBI Manual for Evidence Synthesis*. JBI; 2024. Available from: https://synthesismanual.jbi.global.https://doi.org/10.46658/JBIMES-24-01
6. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), Learning analytics: From research to practice (pp. 61–75). Springer.
7. Ball-Rokeach, S. J., & DeFleur, M. L. (1976). A dependency model of mass-media effects. Communication Research, 3(1), 3–21. https://doi.org/10.1177/009365027600300101
8. Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Prentice-Hall.
9. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21) (pp. 610–623). https://doi.org/10.1145/3442188.3445922
10. Blease, C., Kaptchuk, T. J., Bernstein, M. H., Mandl, K. D., Halamka, J. D., & DesRoches, C. M. (2019). Artificial intelligence and the future of primary care: Exploratory qualitative study of UK general practitioners' views. Journal of Medical Internet Research, 21(3), e12802. https://doi.org/10.2196/12802
11. Bonezzi, A., Ostinelli, M., & Melzner, J. (2022). Do we understand humans better than algorithms? Journal of Experimental Psychology: General, 151(8), 1916- 1933. https://doi.org/10.1037/xge0001181
12. Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. Annual Review of Psychology, 68, 627–652. https://doi.org/10.1146/annurev-psych-010416-043958
13. Calvo, R. A., & Peters, D. (2014). Positive computing: Technology for wellbeing and human potential. MIT Press.
14. Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. Journal of Marketing Research, 56(5), 809–825. https://doi.org/10.1177/0022243719851788
15. Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. Procedia Computer Science, 136, 16–24.
16. Chedrawi, C., Haddad, G., Tarhini, A., Osta, S., & Kazoun, N. (2025). Exploring the impact of responsible AI usage on users' behavioral intention. Journal of Innovation & Knowledge, 10(1), 100813.
17. Choudhary, S., Kaushik, N., Sivathanu, B., & Rana, N. P. (2024). Assessing factors influencing customers' adoption of AI-based voice assistants. Journal of Computer Information Systems, 65(5), 592–609. https://doi.org/10.1080/08874417.2024.2312858

18. Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7–19. https://doi.org/10.1093/analys/58.1.7
19. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008
20. de Graaf, M. M. A., Ben Allouch, S., & van Dijk, J. A. G. M. (2016). Long-term evaluation of a social robot in real homes. Interaction Studies, 17(3), 461–490.
21. Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. Plenum Press.
22. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126. https://doi.org/10.1037/xge0000033
23. Du, Q., Ren, Y., Meng, Z., He, H., & Meng, S. (2025). The Efficacy of Rule-Based versus Large Language Model-Based chatbots in Alleviating Symptoms of Depression and Anxiety: Systematic Review and Meta-Analysis. Journal of Medical Internet Research, 27, e78186. https://doi.org/10.2196/78186
24. Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy. International Journal of Information Management, 57, Article 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002
25. Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642
26. Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. Journal of Medical Internet Research, 21(5), e13216. https://doi.org/10.2196/13216
27. Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. JMIR Mental Health, 4(2), e19. https://doi.org/10.2196/mental.7785
28. Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. Minds and Machines, 30(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1
29. Folk, D., & Dunn, E. (2026). How does turning to AI for companionship predict loneliness and vice versa? Psychological Science, 37(4), 276–286. https://doi.org/10.1177/09567976261427747
30. Fraiwan, M., & Khasawneh, N. (2023). A review of ChatGPT applications in education, marketing, software engineering, and healthcare: Benefits, drawbacks, and research directions (arXiv:2305.00237). arXiv. https://doi.org/10.48550/arXiv.2305.00237
31. Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. JMIR Mental Health, 5(4), e64. https://doi.org/10.2196/mental.9782
32. Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 32(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1
33. Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. JMIR mHealth and uHealth, 6(11), e12106. https://doi.org/10.2196/12106

34. Kasneci, E., Sessler, K., Kuchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Gunnemann, S., Hullermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274
35. Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., & Singh, I. (2019). Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. Biomedical Informatics Insights, 11, 1178222619829083. https://doi.org/10.1177/1178222619829083
36. Kundu, A., & Bej, T. (2025). Psychological impacts of AI use on school students: A systematic scoping review of the empirical literature. Research and Practice in Technology Enhanced Learning, 20, 30. https://doi.org/10.58459/rptel.2025.20030
37. Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., & Coiera, E. (2018). Conversational agents in healthcare: A systematic review. Journal of the American Medical Informatics Association, 25(9), 1248–1258. https://doi.org/10.1093/jamia/ocy072
38. Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H. C., Paulus, M. P., Krystal, J. H., & Jeste, D. V. (2021). Artificial Intelligence for Mental Health Care: Clinical Applications, Barriers, Facilitators, and Artificial Wisdom. Biological psychiatry. Cognitive neuroscience and neuroimaging, 6(9), 856–864. https://doi.org/10.1016/j.bpsc.2021.02.001
39. Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90–103. https://psycnet.apa.org/record/2019-16202-008
40. Luckin, R. (2018). Machine learning and human intelligence: The future of education for the 21st century. UCL Institute of Education Press.
41. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson.
42. Luo, X., Wang, Z., Tilley, J. L., Balarajan, S., Bassey, U.-A., & Cheang, C. I. (2025). Seeking emotional and mental health support from generative AI: Mixed-methods study of ChatGPT user experiences. JMIR Mental Health, 12, e77951. https://doi.org/10.2196/77951
43. Meshi, D., & Ellithorpe, M. E. (2021). Problematic social media use and social support received in real-life versus on social media: Associations with depression, anxiety and social isolation. Addictive Behaviors, 119, 106949. https://doi.org/10.1016/j.addbeh.2021.106949
44. Miner, A. S., Shah, N., Bullock, K. D., Arnow, B. A., Bailenson, J., & Hancock, J. (2019). Key considerations for incorporating conversational AI in psychotherapy. Frontiers in Psychiatry, 10, 746. https://doi.org/10.3389/fpsyt.2019.00746
45. Montag, C., & Hegelich, S. (2020c). Understanding detrimental aspects of social media use: Will the real culprits please stand up? Frontiers in Sociology, 5, 599270. https://doi.org/10.3389/fsoc.2020.599270
46. Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the Implementation of ChatGPT in Education: A Systematic Review. Computers, 12(8), 153. https://doi.org/10.3390/computers12080153
47. Moussawi, S., Koufaris, M., & Benbunan-Fich, R. (2020). How perceptions of intelligence and anthropomorphism affect adoption of personal intelligent agents. Electronic Markets, 31(2), 343–364. https://doi.org/10.1007/s12525-020-00411-w
48. Ni, W., & Li, J. (2026). Mapping the impact of generative AI in higher education: a scoping review of psychological and equity dimensions. Frontiers in Psychology, 17, 1856854. https://doi.org/10.3389/fpsyg.2026.1856854
49. Pillai, R., & Sivathanu, B. (2020). Adoption of AI-based chatbots for hospitality and tourism. International Journal of Contemporary Hospitality Management, 32(10), 3199–3226. https://doi.org/10.1108/IJCHM-04-2020-0259

50. Reeves, B., & Nass, C. (1996). The media equation: How people treat computers, television and new media like real people and places. Cambridge University Press.
51. Russell, S. (2019). Human compatible: Artificial intelligence and the problem of control. Viking.
52. Shalu, Verma, N., Dev, K., Bhardwaj, A. B., & Kumar, K. (2025). The Cognitive Cost of AI: How AI anxiety and attitudes influence decision fatigue in daily technology use. Annals of Neurosciences, 33(1), 73–84. https://doi.org/10.1177/09727531251359872
53. Shneiderman, B. (2022). Human-centered AI: Ensuring human control while increasing automation. Oxford University Press.
54. Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google Effects on memory: Cognitive consequences of having information at our fingertips. Science, 333(6043), 776–778. https://doi.org/10.1126/science.1207745
55. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
56. Tadimalla, S. Y., & Maher, M. L. (2024, October). AI literacy for all: Adjustable interdisciplinary socio-technical curriculum. In 2024 IEEE frontiers in education conference (FIE) (pp. 1-9). IEEE.
57. Tuomi, I. (2018). The impact of artificial intelligence on learning, teaching, and education: Policies for the future (EUR 29442 EN). Publications Office of the European Union. https://doi.org/10.2760/12297
58. Turkle, S. (2011). Alone together: Why we expect more from technology and less from each other. Basic Books.
59. Tzirides, A. O. (O.), Saini, A., Zapata, G. C., Searsmith, D., Cope, B., Kalantzis, M., Castro, V., Kourkoulou, T., Jones, J. W., Abrantes da Silva, R., Whiting, J. K., & Kastania, N. P. (2023). Generative AI: Implications and applications for education. arXiv. https://doi.org/10.48550/arXiv.2305.07605
60. Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and Conversational Agents in Mental Health: A review of the Psychiatric landscape. The Canadian Journal of Psychiatry, 64(7), 456–464. https://doi.org/10.1177/0706743719828977
61. Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. Learning, Media and Technology, 45(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995
62. Woolf, B. P. (2010). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Elsevier.
63. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic Review of Research on Artificial intelligence Applications in Higher Education – Where are the educators? International Journal of Educational Technology in Higher Education, 16(1). https://doi.org/10.1186/s41239-019-0171-0
64. Zhai, X. (2023). ChatGPT user experience: Implications for education. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4312418
65. Zhang, C., & Magerko, B. (2025). Generative AI Literacy: A Comprehensive framework for literacy and responsible use. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2504.19038