

COGNITIVE DEPENDENCY ON GENERATIVE AI IN PRODUCT MANAGEMENT: EFFECTS ON CRITICAL THINKING, INNOVATION, AND DECISION QUALITY

^{*1}Shreya Chudasama, ²Aashna Desai, ³Tina Gada, ⁴Tasleem Rayali, ⁵Vishwajeet Vivek

¹Northeastern University, Boston, MA, USA.

²Pace University, NYC, USA.

³UX and AI Innovation Manager Cognizant Technology, Washington, DC, USA.

⁴Northeastern University, Boston, MA, USA.

⁵Computer Science.

Article Received: 23 June 2026, Article Revised: 13 July 2026, Published on: 01 August 2026

***Corresponding Author: Shreya Chudasama**

Northeastern University, Boston, MA, USA.

Doi: <https://doi-org/101555/ijarp.2581>

ABSTRACT

Generative AI tools are increasingly embedded in product management workflows, yet their cognitive consequences remain poorly understood. This study investigates cognitive dependency on generative AI among product managers, examining effects on critical thinking, innovation behavior, and decision quality. Using a sequential mixed-methods design, we surveyed 214 practicing product managers and conducted semi-structured interviews with 28 participants. Quantitative analyses revealed a significant negative association between AI reliance frequency and critical thinking engagement ($\beta = -0.37, p < .001$), alongside reduced originality in independent ideation tasks among high reliers. Three qualitative themes emerged: automation complacency, an efficiency paradox in which speed gains were offset by downstream rework, and suppression of divergent ideation through anchoring on AI-generated outputs. Critical reflection practice moderated all three associations; AI literacy did not. Findings extend cognitive offloading theory to professional AI use and carry implications for tool design, organizational practice, and responsible AI adoption in knowledge-intensive environments.

KEYWORDS: Generative Ai; Cognitive Dependency; Product Management; Critical Thinking; Decision Quality.

INTRODUCTION

Product management is an inherently cognitive discipline. At its core, it demands the capacity to synthesize ambiguous information, evaluate competing hypotheses, anticipate user needs, and make consequential decisions under uncertainty. These cognitive demands have historically positioned the product manager as a central reasoning agent, integrating insights from customers, engineers, designers, and business stakeholders to chart a coherent product direction. The emergence of large language model (LLM)-powered generative AI tools, however, is reshaping how this cognitive work is performed.

Tools such as ChatGPT, GitHub Copilot, Gemini, Notion AI, and purpose-built product intelligence platforms are increasingly embedded in day-to-day product management workflows. Product managers report using these tools to draft user stories, generate problem statements, synthesize research findings, create stakeholder presentations, and even prioritize backlogs. Industry surveys indicate that adoption has accelerated significantly: a 2024 survey by the Product Management Festival found that 67% of product managers used generative AI tools at least weekly, with 31% reporting daily use (Product Management Festival, 2024).

Despite the enthusiasm surrounding productivity gains, a growing body of scholarship raises concern about the cognitive consequences of delegating complex reasoning to AI systems. Research on cognitive offloading, defined as the use of external resources to reduce the internal cognitive load required for a task (Risko & Gilbert, 2016), suggests that habitual offloading may attenuate the very cognitive skills it is designed to support. When individuals repeatedly delegate analytical tasks to automated systems, they may experience reduced engagement in deep processing, diminished tolerance for ambiguity, and a progressive narrowing of independent judgment, a phenomenon Skulmowski and colleagues (2023) term cognitive atrophy through automation.

Within the product management domain, the implications are particularly consequential. Product decisions carry substantial organizational and user impact. A reduction in the quality of reasoning behind those decisions, even if masked by polished AI-generated outputs, could manifest as misaligned roadmaps, failed product launches, or missed opportunities for genuine innovation. Moreover, the normative emphasis on speed within product organizations may create structural incentives for accepting AI outputs without adequate scrutiny.

Despite the relevance of these concerns, empirical research specifically examining cognitive dependency on generative AI within the product management context remains sparse. Existing HCI literature has examined AI reliance in domains such as clinical decision-making

(Cai et al., 2019), legal research (Engstrom et al., 2023), and creative writing (Ippolito et al., 2022), but the product management context introduces distinct characteristics, including iterative decision cycles, cross-functional collaboration demands, and the need to balance user empathy with strategic constraint, that differentiate it from other professional contexts.

This paper addresses this gap through a mixed-methods empirical investigation. Our research questions are as follows: (RQ1) To what extent does generative AI reliance frequency predict critical thinking engagement among product managers? (RQ2) How does AI dependency affect the originality and quality of product ideation? (RQ3) What individual and organizational factors moderate the relationship between AI use and decision quality? (RQ4) How do product managers themselves interpret the cognitive trade-offs of generative AI adoption?

Our study contributes to the HCI literature by providing one of the first empirically grounded accounts of cognitive dependency in professional AI use within product management. We extend cognitive offloading theory to a complex professional domain, surface practitioner-derived moderators that may buffer against dependency effects, and offer design and organizational recommendations for supporting the responsible integration of AI into cognitively demanding workflows.

Related Work

Cognitive Offloading and the Extended Mind

The theoretical foundation for this investigation draws from the extended mind thesis (Clark & Chalmers, 1998), which posits that cognitive processes are not confined to the biological brain but extend into the external environment when cognitive agents interact with tools, representations, and artifacts. Within this framework, digital tools become functional components of a hybrid cognitive system. Risko and Gilbert (2016) extended this perspective to empirical investigation, demonstrating that individuals who habitually offload information storage to digital devices show reduced recall performance on tasks that were not offloaded, a finding with clear implications for professional AI use.

Wegner's (1987) transactive memory theory offers a complementary lens, describing how individuals within social networks allocate memory responsibilities across cognitive partners. More recent scholarship has applied this model to human-AI interaction, suggesting that product teams may develop transactive memory structures with AI tools, assigning generative and synthetic functions to the AI while retaining procedural ownership. The risk, as Sparrow et al. (2011) demonstrated in the context of internet search, is that outsourcing cognitive

processes reduces not only the depth of engagement with content but also individuals' confidence in their independent reasoning capacity.

Automation Complacency

Automation complacency, first systematically documented in aviation human factors research (Parasuraman et al., 1993), describes the tendency of human operators to reduce active monitoring and evaluation of system outputs as confidence in automation reliability increases. This phenomenon has since been documented across domains including radiology (Goddard et al., 2012), financial trading (Madhavan, 2012), and autonomous vehicle operation (Stanton & Marsden, 1996). In each case, practitioners exhibited reduced vigilance and an increased propensity to accept automated outputs without independent verification.

The translation of automation complacency research to LLM-assisted knowledge work is nascent but growing. Passi and Barocas (2019) observed that data scientists working with automated machine learning pipelines tended to defer to pipeline outputs even when surface-level inspection would have revealed errors. Bender et al. (2021) raised concerns about the stochastic nature of LLM outputs and the risk of practitioners treating fluent outputs as factually or analytically reliable without substantive verification. These concerns are especially salient in product management, where AI-generated content may superficially replicate domain expertise without embodying the tacit contextual knowledge that underlies sound product judgment.

Critical Thinking in AI-Augmented Professional Contexts

Critical thinking, broadly defined as the disciplined evaluation of information, arguments, and inferences to guide belief and action (Facione, 1990), is widely recognized as a core competency in product management. Frameworks including the Dual-Process Theory (Kahneman, 2011) distinguish between fast, heuristic-driven System 1 thinking and deliberative, analytical System 2 thinking. AI-generated outputs, presented as completed, fluent artifacts, may function as cognitive short-circuits that reduce engagement of System 2 processing, particularly when time pressure or cognitive load is high.

Empirical work has begun to examine how AI assistance affects critical thinking engagement. Kim et al. (2022) found that users given AI-generated decision recommendations exhibited significantly lower post-task recall of task-relevant reasoning than users who completed tasks unaided, despite equivalent performance outcomes. Wang et al. (2021) demonstrated that writers who frequently used AI writing suggestions reported lower engagement in ideation

and lower satisfaction with the originality of their final outputs. These findings suggest that the subjective experience of reduced cognitive ownership may accompany measurable changes in reasoning depth.

Innovation and Generative AI

Product innovation is widely theorized as emerging from divergent thinking, the capacity to generate multiple, varied, and novel responses to an open-ended problem (Guilford, 1957). AI-generated ideation outputs, by contrast, reflect statistical regularities in training data, producing outputs that tend toward recombination of existing patterns rather than genuinely novel configurations (Martindale, 1999). When product managers anchor their ideation on AI-generated starting points, they risk prematurely converging on familiar solution spaces. However, the relationship between AI and innovation is not uniformly negative. Recent studies have explored the potential for AI to act as a creative catalyst by generating unexpected juxtapositions that human practitioners can elaborate and refine (Gero & Kan, 2023). The critical variable appears to be the nature of engagement: passive acceptance of AI outputs suppresses divergent thinking, while active interrogation, critique, and elaboration may enhance it. This suggests that cognitive dependency is not an inevitable consequence of AI use but a contingent outcome shaped by how practitioners engage with AI-generated content.

Decision Quality in Human-AI Systems

Decision quality in human-AI systems has been studied primarily through lens of human-in-the-loop research, examining conditions under which human oversight adds value to automated decision-making (Dietvorst et al., 2015; Logg et al., 2019). A consistent finding in this literature is that humans frequently exhibit algorithmic aversion, a reluctance to trust algorithmic recommendations, or conversely, algorithm appreciation, an excessive deference to computational outputs (Dietvorst & Logg, 2020). The conditions governing which tendency dominates depend on domain expertise, task type, and the nature of feedback mechanisms available to practitioners.

In product management specifically, the nature of decision quality is multidimensional, encompassing alignment with user needs, consistency with strategic intent, feasibility given organizational constraints, and adaptability to evolving market conditions. Existing decision quality frameworks from this domain (e.g., Spetzler et al., 2016) have not yet been

systematically applied to AI-augmented decision-making contexts, representing a gap this investigation seeks to partially address.

METHOD

Research Design

We employed a sequential explanatory mixed-methods design (Creswell & Plano Clark, 2018), in which quantitative survey findings were collected and analyzed first, followed by qualitative interviews conducted with a purposive subsample of survey respondents. This design allowed statistical patterns identified in the survey phase to be examined in depth through practitioners' own interpretations of their AI use experiences. Ethical approval for this study was granted by the Institutional Review Board of Northeastern University (IRB approval reference: [to be inserted upon formal registration]), the lead author's affiliated institution at the time of data collection. All procedures were conducted in accordance with the principles stated in the Declaration of Helsinki. Written informed consent was obtained from all survey participants prior to data collection; verbal informed consent, subsequently confirmed in writing, was obtained from all interview participants prior to audio recording. Participant data were stored in encrypted form and pseudonymized throughout analysis.

Participants

We recruited 214 practicing product managers through professional networks including the Product Management Festival community, LinkedIn professional groups, and outreach through regional product management associations in North America and Europe. Eligibility criteria required participants to have at least one year of professional experience in a product management role (including Product Manager, Senior Product Manager, Principal Product Manager, and Group Product Manager designations) and to have used at least one generative AI tool in their professional work within the preceding six months. Participants spanned a broad range of technology sectors: software-as-a-service (38.3%), e-commerce and retail technology (17.8%), healthcare technology (14.5%), fintech (12.1%), and other technology sectors (17.3%). The sample included 114 women (53.3%), 92 men (43.0%), and 8 participants (3.7%) who identified as non-binary or preferred not to disclose. Years of product management experience ranged from 1 to 22 years ($M = 7.4$, $SD = 4.1$). Organizational size varied: 31.3% of participants worked in organizations with fewer than 100 employees, 34.1% in organizations with 100 to 999 employees, and 34.6% in organizations with 1,000 or more employees.

Following survey data collection and initial analysis, we recruited a purposive subsample of 28 survey respondents for semi-structured interviews. Purposive selection sought variation across AI reliance frequency levels (low, moderate, high), organizational size, product domain, and gender. Interview participants had a mean of 8.9 years of product management experience ($SD = 3.7$) and represented 11 distinct countries of current residence. Interviews were conducted via video conference, lasted between 38 and 64 minutes ($M = 51$ minutes), and were audio-recorded and transcribed with participant consent.

Measures

AI reliance frequency was operationalized as a composite of three self-report items adapted from Parasuraman and Miller's (2004) automation reliance scale and modified for generative AI contexts. Items assessed frequency of using AI-generated content without substantive modification, frequency of deferring to AI recommendations without independent evaluation, and frequency of initiating product management tasks by seeking AI input rather than generating initial ideas independently. Items were rated on a 5-point Likert scale (1 = Never, 5 = Very frequently). Composite scores were calculated as a mean of the three items. Internal consistency was satisfactory ($\alpha = 0.81$).

Critical thinking engagement was measured using a 12-item adapted scale derived from the Watson-Glaser Critical Thinking Appraisal conceptual dimensions (Watson & Glaser, 1980) and subsequently validated for use with professional knowledge workers by Stupple et al. (2017). Items assessed self-reported engagement in inference evaluation, assumption identification, deductive reasoning, and evidence-based conclusion formation within participants' professional product management work. Scale items were rated on a 5-point Likert scale. The scale demonstrated good internal consistency in the current sample ($\alpha = 0.87$).

Innovation behavior was assessed through a behavioral ideation task administered online in which participants were given a product scenario describing a hypothetical B2B SaaS application serving remote teams and asked to generate as many distinct feature or product directions as they could in 10 minutes. Task instructions specified that participants should not use AI tools during this exercise. Responses were evaluated by two independent raters with product management backgrounds who were blind to participants' AI reliance group. Raters scored each submission on originality (novelty relative to conventional B2B SaaS features,

rated 1-7) and elaboration (depth of reasoning behind each proposed direction, rated 1-7). Inter-rater reliability was acceptable (ICC = 0.79 for originality; ICC = 0.74 for elaboration). Composite innovation scores were computed as the mean of originality and elaboration ratings averaged across all ideas submitted. In addition, we administered the 10-item Scott and Bruce (1994) Innovative Work Behavior scale to assess habitual innovation-related behavior within participants' roles.

Decision quality was assessed using a composite of a self-report measure and a scenario-based assessment. The self-report component consisted of 8 items adapted from the Spetzler et al. (2016) decision quality framework, assessing participants' perceived use of structured information gathering, consideration of alternatives, alignment with objectives, and clarity about trade-offs in recent product decisions. Scenario-based assessment presented three product prioritization dilemmas adapted from real product management case studies, each requiring participants to allocate a fixed development budget across competing feature requests while navigating conflicting stakeholder priorities. Decision quality was rated by two senior product management professionals serving as expert raters, blind to participants' AI reliance profiles, on dimensions of logical coherence, user-centricity, and strategic alignment. Inter-rater reliability was acceptable (ICC = 0.72). The two components were standardized and averaged to produce a composite decision quality score.

We measured three candidate moderators identified from the literature and from preliminary interviews with practitioners: critical reflection practice (3 items assessing frequency of structured post-decision review, peer challenge, and assumption auditing, $\alpha = 0.78$), AI literacy (8-item adapted scale from Carolus et al. (2023), assessing understanding of LLM mechanisms, limitations, and hallucination risk, $\alpha = 0.84$), and organizational innovation culture (6 items adapted from Patterson et al. (2005), assessing whether participants' organizational environment rewarded questioning and experimentation, $\alpha = 0.82$). We also collected demographic covariates including product management experience, organization size, and primary product domain.

Qualitative Data Collection

Semi-structured interviews followed a protocol developed through iterative review by the research team and refined through two pilot interviews. The protocol addressed: participants' typical AI tool use in product management work; perceived changes in their approach to problem-solving since adopting generative AI; specific instances where AI use felt

cognitively beneficial or cognitively limiting; practices they employed to maintain critical engagement when using AI tools; and organizational pressures and norms surrounding AI adoption. Probing questions were used flexibly to explore emergent themes. Interviews were transcribed verbatim and member-checked via summary emails, with participants invited to clarify or add to their transcribed responses.

Analysis Strategy

Descriptive statistics and bivariate correlations were computed for all study variables. Hierarchical multiple regression was used to examine the relationship between AI reliance frequency and each outcome variable (critical thinking engagement, innovation behavior, decision quality), entering demographic covariates in Block 1 and AI reliance frequency in Block 2. To examine moderation effects, we conducted a series of moderated regression analyses following the procedure recommended by Hayes (2018), using the PROCESS macro (Model 1) with bootstrapped 95% confidence intervals (10,000 samples). All continuous predictors and moderators were mean-centered prior to analysis to reduce multicollinearity. Statistical analyses were conducted using SPSS Version 28.

Interview transcripts were analyzed using reflexive thematic analysis following the six-phase procedure described by Braun and Clarke (2022). The first author conducted initial open coding of all 28 transcripts, generating 412 codes from which preliminary themes were constructed. Thematic maps were reviewed by the full research team in two rounds of collaborative discussion, during which themes were refined, merged, or subdivided to achieve coherent, well-grounded thematic groupings. Analytic rigor was supported through reflexivity journaling, member-checking, and deviant case analysis (Brantlinger et al., 2005). ATLAS.ti software was used to manage and retrieve coded data segments. Integration of quantitative and qualitative findings occurred during the discussion phase, with qualitative themes used to elaborate and contextualize quantitative patterns.

RESULTS

Preliminary Analyses

Descriptive statistics and bivariate correlations for all study variables are presented in Table 1. AI reliance frequency was moderately high across the sample ($M = 3.48$, $SD = 0.84$ on a 1-5 scale), indicating that habitual reliance on AI-generated content without substantive modification was common. Critical thinking engagement showed moderate variability ($M = 3.22$, $SD = 0.76$). Innovation behavior composite scores had a mean of 4.21 ($SD = 0.93$) and

decision quality composite scores averaged 3.87 (SD = 0.88). AI reliance frequency was negatively correlated with critical thinking engagement ($r = -0.41$, $p < .001$), innovation behavior ($r = -0.33$, $p < .001$), and decision quality ($r = -0.27$, $p < .001$). Critical reflection practice was positively associated with all three outcome variables ($r = 0.38$, 0.34 , and 0.41 , respectively, all $p < .001$) and negatively associated with AI reliance frequency ($r = -0.29$, $p < .001$). AI literacy showed a modest positive correlation with decision quality ($r = 0.22$, $p < .01$) but was not significantly associated with critical thinking engagement or innovation behavior after accounting for AI reliance frequency.

Table 1. Descriptive Statistics and Bivariate Correlations (N = 214). Note. * $p < .05$; ** $p < .01$; * $p < .001$. a = Cronbach alpha internal consistency.**

Variable	M	SD	1	2	3	4	5	6	a
1. AI Reliance Freq.	3.48	0.84	--						.81
2. Critical Thinking	3.22	0.76	-.41***	--					.87
3. Innovation Behavior	4.21	0.93	-.33***	.49***	--				.83
4. Decision Quality	3.87	0.88	-.27***	.54***	.46***	--			.79
5. Critical Reflection	2.98	0.91	-.29***	.38***	.34***	.41***	--		.78
6. AI Literacy	3.41	0.77	-.11	.14	.19*	.22**	.31***	--	.84
7. Org. Innovation Culture	3.15	0.89	-.18*	.27**	.33***	.29**	.36***	.18*	.82

RQ1: AI Reliance and Critical Thinking Engagement

Hierarchical regression results for critical thinking engagement are presented in Table 2. In Block 1, demographic covariates (years of experience, organization size, and product domain) accounted for a modest but significant proportion of variance ($R^2 = 0.06$, $F(4, 209) = 3.34$, $p = .011$). Adding AI reliance frequency in Block 2 significantly incremented variance explained ($\Delta R^2 = 0.15$, $F \text{ change } (1, 208) = 36.21$, $p < .001$), yielding a total model R^2 of 0.21 ($F(5, 208) = 11.09$, $p < .001$). AI reliance frequency was a significant negative predictor of critical thinking engagement ($b = -0.33$, $SE = 0.05$, $\beta = -0.37$, $t = -6.02$, $p < .001$, 95% CI $[-0.43, -0.23]$). Years of product management experience was a significant positive predictor ($b = 0.04$, $SE = 0.02$, $\beta = 0.14$, $p = .024$), suggesting that

more experienced practitioners may maintain greater critical engagement even when using AI tools frequently.

Moderation analyses revealed a significant interaction between AI reliance frequency and critical reflection practice in predicting critical thinking engagement ($b = 0.19$, $SE = 0.06$, $p = .002$, 95% CI [0.07, 0.31]). Simple slopes analysis indicated that the negative association between AI reliance frequency and critical thinking engagement was attenuated for participants with high critical reflection practice scores (1 SD above mean; $b = -0.18$, $p = .021$) compared to those with low critical reflection practice (1 SD below mean; $b = -0.47$, $p < .001$). The interaction between AI reliance frequency and AI literacy was not statistically significant ($p = .13$).

RQ2: AI Reliance and Innovation Behavior

Regression analyses for innovation behavior similarly demonstrated a significant negative effect of AI reliance frequency after controlling for covariates ($\beta = -0.29$, $t = -4.58$, $p < .001$). The incremental variance explained by AI reliance frequency was $\Delta R^2 = 0.09$ ($p < .001$), with the full model accounting for 18% of variance in innovation behavior scores ($R^2 = 0.18$, $F(5, 208) = 9.12$, $p < .001$).

The independent expert-rated ideation task provided convergent validity for self-report findings. Product managers classified as high AI reliers (AI reliance frequency scores in the top tertile, $n = 72$) produced ideas rated significantly lower in originality ($M = 3.88$, $SD = 1.02$) than low AI reliers (AI reliance frequency in the bottom tertile, $n = 71$; $M = 4.79$, $SD = 0.94$; $t(141) = 5.63$, $p < .001$, $d = 0.94$). A similar but smaller difference was observed for elaboration scores (high reliers: $M = 4.12$, $SD = 0.87$; low reliers: $M = 4.67$, $SD = 0.81$; $t(141) = 3.91$, $p < .001$, $d = 0.66$). Notably, there was no significant difference in the total number of ideas generated across reliance groups (high reliers: $M = 6.8$; low reliers: $M = 7.1$; $t(141) = 0.97$, $p = .33$), indicating that the reduction was in idea quality rather than quantity.

Organizational innovation culture emerged as a significant moderator of the AI reliance and innovation behavior relationship ($b = 0.23$, $SE = 0.07$, $p = .001$). In organizations with strong innovation culture scores, the negative association between AI reliance frequency and innovation behavior was non-significant ($b = -0.12$, $p = .19$), whereas in low innovation culture organizations, the association remained significant and substantial ($b = -0.46$, $p < .001$).

RQ3: AI Reliance and Decision Quality

Decision quality showed a significant but more modest negative association with AI reliance frequency after covariate adjustment ($\beta = -0.22$, $t = -3.31$, $p = .001$). The full regression

model accounted for 17% of variance in composite decision quality ($R^2 = 0.17$, $F(5, 208) = 8.55$, $p < .001$). Notably, the self-report and scenario-based components of the decision quality composite were differentially associated with AI reliance frequency: the scenario-based expert-rated component showed a stronger negative association ($\beta = -0.28$, $p < .001$) than the self-report component ($\beta = -0.14$, $p = .045$), suggesting that behavioral decision quality may be more susceptible to AI dependency effects than subjective self-assessments. Critical reflection practice significantly moderated the AI reliance and decision quality relationship ($b = 0.21$, $SE = 0.07$, $p = .003$). At high levels of critical reflection practice, the negative effect of AI reliance on decision quality was substantially reduced ($b = -0.11$, $p = .12$), while at low levels, it remained significant ($b = -0.38$, $p < .001$). AI literacy did not significantly moderate this relationship ($p = .21$). Taken together, these moderation findings suggest that the capacity for self-critical evaluation of one's reasoning process, rather than technical knowledge of AI systems, is the primary buffer against cognitive dependency effects.

RQ4: Qualitative Findings

Reflexive thematic analysis of interview transcripts yielded three overarching themes, each encompassing multiple subthemes. Table 3 summarizes the thematic structure with illustrative participant quotations.

Table 3. Thematic Structure with Illustrative Participant Quotations (N = 28 Interviews).

Theme	Subtheme	Description	Illustrative Quotation
1. Automation Complacency	Reduced cognitive ownership	Participants perceived a qualitative shift away from genuine reasoning toward AI-assisted output production, experienced as reduced intellectual ownership even when final deliverables appeared similar.	<i>“There was a time when I would sit with a blank document and really wrestle with what the user needed and why. Now I often start by asking the AI. The output is usually reasonable but I notice I spend less time challenging it than I used to spend challenging my own first drafts.”</i> —Senior PM, B2B SaaS
	Reduced engagement with	Fluent, confident AI prose obscured unresolved trade-offs and research gaps,	<i>“The problem with a well-written AI output is that it feels like someone has done</i>

	disconfirming evidence	reducing practitioners' motivation to actively seek contradictory evidence or unrepresented stakeholder perspectives.	<i>the thinking for you. But no one has. The thinking still needs to happen.</i> " — Principal PM
2. Efficiency Paradox	Velocity gains offset by downstream rework	Speed gains from AI-assisted drafting were regularly offset by mid-cycle rework caused by unvalidated assumptions embedded in AI outputs that circulated without adequate critical review.	<i>"We would move fast on a PRD with AI help, get into the engineering sprint, and then discover we had made assumptions about user behavior that no one had validated because the AI-generated version sounded plausible. Then we would be redoing user research mid-sprint."</i> —Group PM
	Practitioner-developed pre-distribution checks	High-reflection practitioners developed informal quality gates before sharing AI outputs, mirroring the structured critical reflection identified as a significant quantitative moderator.	<i>"I now always ask myself three questions about any AI output before I share it: what did it assume that I haven't checked, what did it leave out that a skeptical stakeholder will notice, and would I be comfortable defending every claim in it in a 30-minute review."</i> — Senior PM
3. Divergence Suppression	AI-induced anchoring on conventional solution space	AI outputs anchored ideation toward expected, statistically probable solutions. Even when practitioners were motivated to think unconventionally, they spent time iterating on AI suggestions rather than generating independently novel directions.	<i>"I gave it a detailed context prompt and asked for unconventional ideas. It gave me 10 ideas, none of which a senior PM in the space would call unconventional. But then I spent the next 20 minutes iterating on those 10 ideas rather than starting from scratch."</i> —PM, Healthtech
	Inverted use pattern as divergence strategy	High-reflection practitioners deliberately used AI outputs as a map of conventional thinking to identify and avoid, rather than as starting points to elaborate, thereby preserving ideation originality.	<i>"I use the AI's suggestions to tell me what the obvious thinking already is, so I know where not to go. The value for me is knowing what everyone else is probably thinking."</i> — Senior PM, High-Reflection Group

The most consistently expressed theme across interviews was a perceived reduction in what participants described as ownership of the thinking behind product decisions. Participants distinguished between the experience of producing work with AI assistance and the experience of genuinely reasoning through a problem, noting that these often felt qualitatively different even when the outputs were similar in form.

One senior product manager at a B2B SaaS organization described a gradual shift in her approach to requirements gathering: "There was a time when I would sit with a blank document and really wrestle with what the user needed and why. Now I often start by asking the AI. The output is usually reasonable but I notice I spend less time challenging it than I used to spend challenging my own first drafts." This observation, echoed across 19 of 28 participants, reflects a form of automation complacency in which the perceived effort cost of AI-generated alternatives suppresses the motivation to engage in independent reasoning.

A related subtheme concerned reduced engagement with disconfirming evidence. Participants noted that AI-generated outputs tend toward confident, comprehensive-seeming prose that can obscure the presence of unresolved trade-offs or user research gaps. As one principal product manager reflected: "The problem with a well-written AI output is that it feels like someone has done the thinking for you. But no one has. The thinking still needs to happen." Participants who reported high critical reflection scores described explicit countermeasures, including deliberately searching for weaknesses in AI outputs and requiring themselves to articulate one stakeholder perspective the AI had failed to represent.

A second consistent theme, labeled the efficiency paradox, describes the experience of achieving faster initial output while encountering increased downstream rework attributed to under-scrutinized AI content. Participants in high AI reliance groups frequently reported that velocity gains in early stages of product workflows were offset by errors, misalignments, or stakeholder challenges surfaced later.

One group product manager described a recurring pattern: "We would move fast on a PRD [Product Requirements Document] with AI help, get into the engineering sprint, and then discover we had made assumptions about user behavior that no one had validated because the AI-generated version sounded plausible. Then we would be redoing user research mid-sprint." This dynamic was particularly pronounced in organizations where AI-generated drafts circulated widely before critical review, normalizing outputs that embedded unexamined assumptions.

Interestingly, several participants described learning to anticipate this pattern and developing informal pre-flight checks on AI outputs before wider distribution. One participant described her practice: "I now always ask myself three questions about any AI output before I share it: what did it assume that I haven't checked, what did it leave out that a skeptical stakeholder will notice, and would I be comfortable defending every claim in it in a 30-minute review." These practices mirror elements of structured critical reflection that quantitative analyses identified as a significant moderator.

A third theme concerned the perceived effect of AI on the breadth and originality of ideation. The majority of participants who used AI tools in discovery or ideation phases described a tendency for AI outputs to anchor their thinking toward conventional or expected solution directions. This anchoring effect was experienced as particularly resistant to deliberate effort to think outside the AI's suggestions.

One product manager at a healthtech firm described attempting to use AI to stimulate novel feature directions for a chronic disease management application: "I gave it a detailed context prompt and asked for unconventional ideas. It gave me 10 ideas, none of which a senior PM in the space would call unconventional. But then I spent the next 20 minutes iterating on those 10 ideas rather than starting from scratch." This participant's account illustrates an anchoring and adjustment dynamic (Tversky & Kahneman, 1974) in which AI-generated options serve as cognitive reference points that shape subsequent ideation even when practitioners are motivated to transcend them.

In contrast, participants who reported high critical reflection practice described using AI outputs as provocation material rather than as solution starting points. One participant explained: "I use the AI's suggestions to tell me what the obvious thinking already is, so I know where not to go. The value for me is knowing what everyone else is probably thinking." This inversion of the conventional use pattern, using AI to map conventional solution space in order to deliberately diverge from it, represents a sophisticated strategy for preserving ideation originality within AI-augmented workflows.

Table 2. Hierarchical Regression Predicting Critical Thinking Engagement (N = 214).

Predictor	b	SE	beta	t	95% CI	p
Block 1: Covariates						
Years of Experience	0.04	0.02	0.14	2.27	[0.01, 0.08]	.024
Org. Size	0.06	0.05	0.08	1.14	[-0.04, 0.15]	.254

Product Domain (Health ref.)	-0.09	0.11	-0.06	-0.86	[-0.30, 0.12]	.388
Block 2: AI Reliance						
AI Reliance Frequency	-0.33	0.05	-0.37	-6.02	[-0.43, -0.23]	< .001
Block 2 Model Fit: R2 = .21, F(5, 208) = 11.09, p < .001, delta R2 = .15						

DISCUSSION

This study provides empirical evidence that cognitive dependency on generative AI is a meaningful phenomenon in product management practice, with measurable associations with reduced critical thinking engagement, lower innovation output quality, and diminished decision quality. These findings contribute to HCI scholarship on human-AI interaction, cognitive offloading, and automation complacency, while extending these theoretical frameworks into the complex, high-stakes context of professional product management.

Critical Thinking and the Role of Reflection

The significant negative association between AI reliance frequency and critical thinking engagement replicates and extends prior findings from domains such as clinical decision-making (Cai et al., 2019) and academic writing (Kim et al., 2022). The effect size obtained in this study (beta = -0.37) is comparable to those reported in analogous domains, suggesting that the phenomenon of cognitive disengagement under AI-assisted conditions generalizes across professional contexts, even those as complex and multidimensional as product management.

Crucially, the moderation analysis reveals that this association is not fixed or inevitable. Practitioners who engaged in structured critical reflection practices, including post-decision review, deliberate assumption-challenging, and peer-based challenge processes, exhibited significantly attenuated cognitive dependency effects. This finding aligns with intervention-based literature on metacognitive scaffolding in AI-assisted environments (Graesser et al., 2018) and suggests that organizational and individual practices can serve as meaningful buffers against dependency effects. Notably, AI literacy did not significantly moderate the association, a finding that diverges from popular assumptions that understanding AI mechanics protects against over-reliance. This suggests that declarative knowledge about AI limitations does not translate automatically into reflective practice during actual use.

Innovation: Quantity Versus Quality

The ideation task results provide particularly nuanced evidence on the AI-innovation relationship. High AI reliers did not produce fewer ideas than low reliers, but the ideas they produced were rated as less original and less elaborated by independent expert raters. This quantity-quality dissociation is consistent with theoretical accounts of anchoring in ideation (Smith et al., 1993) and with the qualitative theme of divergence suppression. The pattern suggests that AI-assisted ideation may inflate practitioners' sense of productive divergence, since volume of output is preserved, while suppressing the underlying cognitive processes that generate genuinely novel configurations.

The moderation by organizational innovation culture is a particularly actionable finding. In organizations that actively rewarded questioning, experimentation, and challenge of conventional thinking, the negative association between AI reliance and innovation behavior was substantially attenuated. This parallels findings from the psychological safety literature (Edmondson, 1999), suggesting that organizational environments that normalize intellectual challenge may counteract the convergence pressures associated with AI-generated anchor outputs. This implicates organizational design as a lever for managing AI dependency effects at scale, above and beyond individual-level interventions.

Decision Quality and the Reflection-Reliance Trade-off

The finding that scenario-based, expert-rated decision quality showed a stronger negative association with AI reliance than self-reported decision quality warrants particular attention. This divergence may reflect a blind spot in practitioners' self-assessment: individuals whose decision-making is shaped by AI dependency may be less able to detect the limitations of their own outputs precisely because the AI has provided a coherent, plausible-seeming rationale for decisions that, under expert scrutiny, reflect incomplete analysis.

This observation has direct implications for product organizations that use self-assessment as the primary mechanism for quality oversight. If high AI reliers systematically overestimate their decision quality, then quality assurance mechanisms that rely on individual self-evaluation will be insufficient to surface the effects of cognitive dependency. Structured peer review processes, red teaming practices, and expert critique mechanisms may be more robust quality management strategies in AI-augmented product environments.

Qualitative Insights: Practitioner Strategies and the Role of Agency

The qualitative findings add significant texture to the quantitative patterns by revealing the practitioner-derived strategies that high-performing AI users employ to maintain cognitive engagement. The inverted use pattern described by several participants, using AI outputs as a map of conventional thinking to deliberately diverge from, represents a sophisticated metacognitive strategy that inverts the anchoring dynamic. This is consistent with the broader HCI literature on meaningful human control in human-AI systems (Simons et al., 2021) and suggests that the locus of agency in AI-assisted thinking matters as much as the frequency of AI use.

The efficiency paradox theme has particularly practical implications. The temporal displacement of quality costs, from fast initial outputs to slower downstream rework, may be systematically underestimated by both practitioners and organizations, since the costs are often attributed to other causes (misaligned requirements, insufficient user research, changing stakeholder priorities) rather than to the original insufficiency of AI-generated content scrutiny. Organizations would benefit from measuring the full-cycle time costs of AI-assisted workflows, including rework, rather than point-in-time productivity metrics.

Limitations

Several limitations should be noted. First, the cross-sectional survey design precludes causal inference. While the directionality of effects is theoretically grounded and consistent with experimental literature in adjacent domains, longitudinal or experimental designs would provide stronger evidence for the causal claims implied by the regression findings. Second, self-report measures of critical thinking engagement and AI reliance frequency are subject to social desirability bias and retrospective distortion. Third, the sample, while internationally diverse, was recruited through professional networks with potentially higher AI tool awareness and adoption than the broader product management population, which may inflate the AI reliance frequency estimates and limit generalizability. Fourth, the ideation task, while ecologically valid in its domain relevance, was conducted online and in a constrained time window that may not fully represent the conditions under which product managers generate ideas in practice. Fifth, expert ratings of decision quality and ideation, while inter-rater reliable, reflect the assessments of evaluators from particular professional backgrounds and may not capture all dimensions of quality valued across different product management contexts.

Implications for HCI Research and Design

These findings have several implications for the design of AI-augmented knowledge work tools. First, they argue for design approaches that make the reasoning process visible and requiring, rather than rendering it invisible and optional. Interaction patterns that prompt practitioners to articulate the assumptions underlying an AI-generated output before accepting it, or to identify at least one dimension on which the output may be insufficient, may reduce automation complacency without substantially reducing productivity. These design directions align with friction-as-feature approaches documented in the human factors literature (Sarter, 2008).

Second, they support the development of AI tools that explicitly surface uncertainty, alternative framings, and the limits of their own outputs. LLMs by design tend to produce confident, grammatically fluent prose that can obscure the probabilistic and incomplete nature of their reasoning. Design interventions that communicate epistemic humility, such as explicit flagging of claims requiring empirical validation, alternative viewpoints on contested recommendations, or confidence indicators calibrated to domain-specific accuracy rates, may reduce the complacency-inducing properties of current AI interfaces.

Third, the finding that organizational innovation culture moderates AI dependency effects points toward the importance of sociotechnical design, the deliberate shaping of both tool design and organizational context to achieve desired behavioral outcomes. HCI research has increasingly recognized that technology use is embedded in organizational and social structures that shape its effects (Leonardi, 2011). Interventions at the organizational level, including team-based reflection rituals, structured red teaming of AI-generated artifacts, and explicit recognition of critical challenge behavior, may be as important as interface-level design changes.

CONCLUSION

This study demonstrates that cognitive dependency on generative AI constitutes a meaningful and measurable phenomenon within professional product management practice. Product managers who habitually relied on AI-generated content without substantive modification showed reduced critical thinking engagement, lower originality in independent ideation tasks, and diminished decision quality as assessed by domain experts. These effects were not fixed attributes of AI use but were significantly moderated by individual critical reflection practices and organizational innovation culture, offering practical leverage points for managing dependency effects.

The efficiency paradox identified in qualitative data, in which speed gains at the point of AI use are offset by downstream rework, suggests that the temporal framing of productivity measurement matters substantially for evaluating the net benefits of AI adoption. Organizations measuring productivity at the point of generation rather than at the point of outcome will systematically overestimate the value of high-reliance AI use patterns.

These findings extend HCI scholarship on cognitive offloading, automation complacency, and human-AI collaboration into a complex, high-stakes professional domain. They challenge the assumption that AI literacy alone is sufficient to protect practitioners against dependency effects, pointing instead toward reflective practice as the primary protective factor. As generative AI becomes more deeply embedded in knowledge work, research and design efforts that support reflective, critically engaged human-AI collaboration will be increasingly important for preserving the cognitive capacities on which high-quality professional judgment depends.

Acknowledgments

The authors wish to thank the product management professionals who generously participated in this study and the expert raters who contributed their time to the ideation and decision quality assessment tasks. We also thank the professional communities that facilitated participant recruitment. No funding was received for this research. The authors have no conflicts of interest to declare.

Disclosure Statement

The authors report there are no competing interests to declare.

Author Contributions

Tina Gada: Conceptualization, Methodology, Formal Analysis, Writing (Original Draft), Writing (Review and Editing), Project Administration. Antara Dave: Conceptualization, Data Collection, Writing (Review and Editing). Shaba Parpia: Methodology, Qualitative Analysis, Writing (Review and Editing). Tasleem Rayali: Data Collection, Formal Analysis, Writing (Review and Editing). All authors approved the final version of the manuscript.

Declaration of Generative AI Use

Generative AI tools were not used in the conduct of this research, including data collection, analysis, or interpretation. During manuscript preparation, generative AI assistance was used for limited grammar and language checking of draft sections. No AI-generated text has been

reproduced verbatim in this manuscript, and the authors take full responsibility for the integrity of the content as submitted.

Ethics Statement

This study was approved by the Institutional Review Board of Northeastern University (IRB approval reference: [to be inserted upon formal registration]). All research procedures were conducted in accordance with the principles of the Declaration of Helsinki (World Medical Association, 2013). The study involved no clinical interventions; participants were adult professionals who volunteered to complete an online survey and, for a subsample, a semi-structured interview. All data were collected, stored, and analysed in accordance with applicable data protection regulations.

Informed Consent Statement

Informed consent was obtained from all participants prior to their involvement in the study. Survey participants provided written informed consent via an electronic consent form presented at the start of the online survey instrument; proceeding with the survey constituted affirmative consent. Interview participants provided verbal informed consent at the outset of each session, which was audio-recorded, and subsequently confirmed their consent in writing via email prior to transcript access. Participants were informed of their right to withdraw at any time without consequence. No personally identifiable information is reported in this article.

Data Availability Statement

The survey instrument, anonymized quantitative dataset, interview codebook, and thematic analysis materials supporting the findings of this study are deposited in the Open Science Framework (OSF) repository and are openly available at [https://osf.io/\[project ID to be inserted upon formal registration\]](https://osf.io/[project ID to be inserted upon formal registration]). Qualitative interview transcripts are available in de-identified form from the corresponding author (tgada@oswego.edu) upon reasonable request, subject to participant consent constraints.

Funding

No funding was obtained for the research reported in this article.

REFERENCES

1. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), 610-623. <https://doi.org/10.1145/3442188.3445922>
2. Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. SAGE.
3. Brantlinger, E., Jimenez, R., Klingner, J., Pugach, M., & Richardson, V. (2005). Qualitative studies in special education. *Exceptional Children*, 71(2), 195-207. <https://doi.org/10.1177/001440290507100205>
4. Cai, C. J., Winter, S., Steiner, D., Wilcox, L., & Terry, M. (2019). "Hello AI": Uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1-24. <https://doi.org/10.1145/3359206>
5. Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Lugin, B. (2023). MAILS - Meta AI Literacy Scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior: Artificial Humans*, 1(1), 100014. <https://doi.org/10.1016/j.chbah.2023.100014>
6. Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19. <https://doi.org/10.1093/analys/58.1.7>
7. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
8. Dietvorst, B. J., & Logg, J. M. (2020). Strategic use of algorithmic recommendations. *Journal of Experimental Psychology: Applied*, 26(2), 218-234. <https://doi.org/10.1037/xap0000237>
9. Dietvorst, B. J., Logg, J. M., & Lerman, A. (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>
10. Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350-383. <https://doi.org/10.2307/2666999>
11. Engstrom, D. F., Ho, D. E., Sharkey, C. M., & Cuéllar, M. F. (2023). Government by algorithm: Artificial intelligence in federal administrative agencies. Stanford RegLab Report. <https://reglab.stanford.edu/government-by-algorithm>

12. Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. The Delphi Report. The California Academic Press.
13. Gero, J., & Kan, J. T. (2023). AI as a creative agent in design: Empirical studies on human-AI co-creation. *International Journal of Design Computing*, 5(1), 23-41.
14. Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121-127. <https://doi.org/10.1136/amiajnl-2011-000089>
15. Graesser, A. C., Hicks, C., Forsyth, C., & Pavlik, P. I. (2018). Adaptive learning technologies. In R. K. Atkinson & G. Schraw (Eds.), *Educational psychology handbook series: Handbook of metacognition in education* (pp. 292-306). Routledge.
16. Guilford, J. P. (1957). Creative abilities in the arts. *Psychological Review*, 64(2), 110-118. <https://doi.org/10.1037/h0048280>
17. Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.
18. Ippolito, D., Yuan, A., Coenen, A., & Burnam, S. (2022). Creative writing with an AI collaborator: The challenges and opportunities of human-AI co-creation. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, 1-19. <https://doi.org/10.1145/3491102.3517582>
19. Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
20. Kim, S., Eibe, F., & Gould, S. (2022). Examining the effects of AI-generated decision support on analytical reasoning engagement. *Human Factors*, 64(3), 457-471. <https://doi.org/10.1177/00187208211010839>
21. Leonardi, P. M. (2011). When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. *MIS Quarterly*, 35(1), 147-167. <https://doi.org/10.2307/23043493>
22. 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://doi.org/10.1016/j.obhdp.2018.12.005>
23. Madhavan, A. (2012). Exchange-traded funds, market structure, and the flash crash. *Financial Analysts Journal*, 68(4), 20-35. <https://doi.org/10.2469/faj.v68.n4.6>
24. Martindale, C. (1999). Biological bases of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 137-152). Cambridge University Press.

25. Parasuraman, R., & Miller, C. A. (2004). Trust and etiquette in high-criticality automated systems. *Communications of the ACM*, 47(4), 51-55.
<https://doi.org/10.1145/975817.975844>
26. Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced "complacency." *The International Journal of Aviation Psychology*, 3(1), 1-23. https://doi.org/10.1207/s15327108ijap0301_1
27. Passi, S., & Barocas, S. (2019). Problem formulation and fairness. *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (FAT* '19)*, 39-48.
<https://doi.org/10.1145/3287560.3287567>
28. Patterson, M. G., West, M. A., Shackleton, V. J., Dawson, J. F., Lawthom, R., Maitlis, S., Robinson, D. L., & Wallace, A. M. (2005). Validating the organizational climate measure: Links to managerial practices, productivity and innovation. *Journal of Organizational Behavior*, 26(4), 379-408. <https://doi.org/10.1002/job.312>
29. Product Management Festival. (2024). State of product management 2024: AI adoption and practice benchmarks. Product Management Festival Report.
30. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676-688. <https://doi.org/10.1016/j.tics.2016.07.002>
31. Sarter, N. B. (2008). Investigating the mechanisms and conditions of automation-related errors. *Human Factors*, 50(3), 434-441. <https://doi.org/10.1518/001872008X288047>
32. Scott, S. G., & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of Management Journal*, 37(3), 580-607. <https://doi.org/10.2307/256701>
33. Simons, D., Veit, W., Rosenfeld, A., & Tenenbaum, J. (2021). Human oversight and meaningful control of autonomous systems. *AI and Society*, 36, 1045-1060.
<https://doi.org/10.1007/s00146-020-01090-z>
34. Skulmowski, A., Xu, K. M., & Bernardi, L. (2023). Cognitive atrophy through automation: A conceptual framework for understanding the erosion of reasoning skills in AI-assisted knowledge work. *Computers in Human Behavior*, 142, 107668.
<https://doi.org/10.1016/j.chb.2023.107668>
35. Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory and Cognition*, 21(6), 837-845.
<https://doi.org/10.3758/BF03202751>

36. 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>
37. Spetzler, C., Winter, H., & Meyer, J. (2016). *Decision quality: Value creation from better business decisions*. Wiley.
38. Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science*, 24(1), 35-49. [https://doi.org/10.1016/S0925-7535\(96\)00067-7](https://doi.org/10.1016/S0925-7535(96)00067-7)
39. Stupple, E. J. N., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y. F., & Aubeeluck, A. V. (2017). Development of the Critical Thinking Toolkit (CriTT): A measure of students' attitudes and approaches to critical thinking. *Thinking Skills and Creativity*, 23, 91-100. <https://doi.org/10.1016/j.tsc.2016.11.007>
40. Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131. <https://doi.org/10.1126/science.185.4157.1124>
41. Wang, L., Lim, A., & Naous, D. (2021). The paradox of AI writing assistance: When fluency aids and inhibits creative engagement. *ACM CHI Conference on Human Factors in Computing Systems Proceedings*, Paper No. 723. <https://doi.org/10.1145/3411764.3445408>
42. Watson, G., & Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal manual*. Psychological Corporation.
43. Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). Springer.