

Expecting Automation: How Users Miss Valuable Experiences in Generative AI

AENEAS STANKOWSKI, German Research Center for AI, Germany

Research on human-AI interaction has documented several phenomena that shape how users engage with generative AI: transactional interface affordances, cognitive offloading tendencies, over-reliance on AI outputs, and organizational pressures favoring rapid completion. These are individually well-established, with parallels in the automation complacency and bias literature. They result in a compounding effect and default orientation toward expecting full automation, or handoff. Drawing on a Research through Design study of AI-supported corporate goal-setting (N=15) and subsequent informal deployment, we observe that this expectation led participants to neglect opportunities for agency and steering, even when offered, creating a barrier between experienced value and chosen engagement. Participants who underwent sustained reflective interaction in structured sessions found it uniquely valuable, yet in informal use quickly attempted to redirect the system toward generating artifacts. We argue that automation expectation, as a convergent effect operating prior to engagement, warrants attention as a design problem in automation experience research, and outline directions for intervention.

CCS Concepts: • **Human-centered computing** → **HCI design and evaluation methods**; **Interaction design**; *Collaborative and social computing*.

Additional Key Words and Phrases: Automation Experience, Generative AI, Human-AI Interaction, Cognitive Outcomes, Knowledge Work, Reflection

ACM Reference Format:

Aeneas Stankowski. 2026. Expecting Automation: How Users Miss Valuable Experiences in Generative AI. In *Proceedings of AutomationXP26 Workshop of the 2026 CHI Conference on Human Factors in Computing Systems, April 14, 2026, Barcelona, Spain*. ACM, New York, NY, USA, 6 pages.

1 Introduction

Generative AI systems have established a dominant interaction pattern: the user states a need, the system produces an artifact, and the exchange concludes [17, 19]. This pattern serves well when the user’s intent is fully formed and the desired output is known. That it can also bypass cognitive processes of value in knowledge work is recognized in the literature on cognitive offloading [12], over-reliance [3], and declining critical thinking under AI assistance [7, 9].

The self-reinforcing characteristics of this effect, and how they align with external incentives and affordances towards a default orientation, has received less attention: users approach generative AI expecting automation, and this expectation forecloses valuable experiences they would value if they occurred. In a Research through Design study of AI-supported goal-setting in a corporate environment, we observed this empirically. Participants who engaged in sustained reflective dialogue with a conversational AI prototype found the experience valuable and expressed strong adoption intent. When the system was made available for informal use, users predominantly attempted to redirect it toward generating artifacts directly, resisting the reflective interaction that evaluation participants had found worthwhile.

Author’s Contact Information: Aeneas Stankowski, German Research Center for AI, Berlin, Germany, aeneas.stankowski@dfki.de.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2026 Copyright held by the owner/author(s).

Manuscript submitted to ACM

This *automation expectation* can be framed as a compound effect and enacted orientation shaped by interface design, learned interaction conventions, and organizational environments. The individual contributing phenomena have been observed in various forms, including in the established literature on automation complacency, bias, and human-agent coordination [8, 10]. Our contribution is in identifying their compound effect in generative AI use, demonstrating its consequence empirically, and arguing that it warrants attention as a design problem in its own right. In what follows, we review documented sources, present the empirical observation, and consider what addressing automation expectation would require.

2 Sources of Automation Expectation

The tendency to approach generative AI with an automation orientation does not arise from a single cause. The contributing phenomena span human-AI interaction, cognitive science, organizational behavior, and the literature on human-automation interaction. We review them here to establish the landscape from which automation expectation emerges.

2.1 Interface Affordances and Interaction Conventions

The dominant interface paradigm for generative AI communicates a transactional model of interaction [17]. Design guidelines emphasize output quality, control over generation parameters, and trust calibration [1], reinforcing the framing of AI as a production tool. Zamfirescu-Pereira et al. [19] documented that non-expert users struggle to engage with AI beyond simple request-response exchanges. Subramonyam et al. [14] identified a “gulf of envisioning” where users have difficulty articulating needs that extend beyond immediate artifact requests. Effective joint activity between humans and agents requires mutual predictability, directability, and maintained common ground [8]. Current generative AI interfaces typically do not invite these; in common use they tend to afford request submission rather than coordinated cognitive activity.

2.2 Cognitive Offloading and Overreliance

The availability of capable AI systems creates incentives for cognitive offloading [12]. Research on AI-assisted decision-making documents overreliance: users accept AI outputs without critical evaluation, even when incorrect [3]. Lee et al. [9] reported self-assessed reductions in critical thinking among knowledge workers using generative AI, linked to decreased confidence in independent reasoning. As users offload more cognitive work, the activation threshold for independent engagement rises. In established automation contexts, trust in reliable systems produces attentional withdrawal from the automated task, reinforced through a positive feedback loop when the consequences of disengagement remain invisible [10].

2.3 Organizational Pressures, Structural Affordances, and Eroding Tasks

Automation expectation is shaped by the environment in which interaction takes place. Organizational affordances [16] structure how tools are understood and used. Time pressure promotes heuristic processing and reliance on defaults [13]. Metric-driven environments deprioritize processes whose value is difficult to quantify [18]. When AI systems are introduced into such environments, organizational incentives align with interface affordances: both point toward rapid completion. Cummings et al. [4] documented how operators in automated environments develop coping strategies for under-stimulating tasks, including task-unrelated thought, distraction, and redirecting engagement toward alternative

activities. Users encountering a reflective interaction that does not match their automation expectation may engage in structurally similar coping: attempting to reshape the interaction into one that does match.

Our formative research in goal-setting surfaced the contextual cues that prime this orientation. The goal-setting process had eroded into a compliance exercise: overloaded with competing functions (development, compensation, steering, evaluation), it provided no signal that cognitive investment would be rewarded. Adverse incentive structures discouraged help-seeking, as admitting uncertainty could signal incompetence. Prior experience with language models had already established artifact generation as the default interaction mode. These cues converge before any interaction with our prototype begins: the task context, the organizational framing of AI as an efficiency tool, and learned interaction conventions collectively signal that automation is expected and appropriate.

2.4 Reduced Choice and Perceived Autonomy

Recent work by Faas et al. [6] demonstrates that AI-supported workflows can restrict perceived choice sets, reducing users' sense of autonomy and work meaningfulness even when decision performance is maintained. When systems present optimized suggestions, users may not perceive alternative modes of engagement as available, further narrowing the interaction toward delegation.

2.5 Moving toward a Composite Account

The sources described above are established findings across different research communities at different levels of abstraction. What we propose is that these sources produce a convergent *automation expectation* in generative AI contexts: interface affordances point toward rapid completion; cognitive tendencies favor offloading; organizational environments reward them.

This framing shares territory with integrated models of automation complacency and bias [10], but while complacency and automation bias develop *during* sustained interaction with reliable automation, through attentional reallocation and trust calibration, automation expectation, as we observe it, is already formed and operates *prior to* engagement. Users in our study did not develop overreliance through repeated experience with the prototype; they arrived with an orientation already formed by converging contextual sources, and this orientation shaped whether certain modes of engagement occurred at all. The compound effect forecloses experience rather than degrading it incrementally. By naming this pre-engagement orientation, we aim to create a design handle: identifying which contributing sources are most active in a given deployment context and taking targeted measures against them.

3 Empirical Observation

We observe two specific interaction tendencies in our Research through Design study [20] in corporate goal-setting at a multinational technology company (detailed in a companion paper under review) comprising contextual inquiry (23 interviews), prototype design, and evaluation with 15 employees in 60-minute structured sessions and follow-up interviews.

The prototype operationalized a principle we termed *cognitive outcomes first*: rather than generating goal statements, the system used structured questioning to prompt articulation of tacit knowledge and generation of new insight. Users selected quality dimensions to work on (e.g., Measurable, Achievable, Adaptable) and the system engaged in persistent inquiry within each before producing any artifact.

3.1 Findings from Structured Evaluation

Across 15 sessions, we observed 164 distinct cognitive outcomes, categorized as Spelling Out (articulation of implicit knowledge; 49.4%) or Cognitive Delta (genuine new insight; 50.6%). The near-equal distribution indicates that the interaction both recovered tacit knowledge and generated understanding that did not previously exist for the participant.

User reception was positive: 73% characterized the persistent questioning as productive rigor; 27% found it adversarial but still acknowledged its cognitive value. All but one (93%) expressed strong intent to use such a system if available. The Adaptable dimension, though selected by only 5 of 15 participants, showed a 92% Cognitive Delta ratio: users who engaged with it generated substantial new insight, but most did not anticipate needing it.

3.2 Findings from Informal Deployment

Following evaluation, the system was made available informally to project stakeholders and to original participants. While initial interest was high, engagement diverged from the evaluation setting. After brief interaction, most users attempted to redirect the dialogue by answering questions with questions, requesting the system to propose goals directly, and resisting providing contextual information. Rather than following the reflective structure, users attempted to bend the interaction toward artifact generation. Original participants showed somewhat more sustained engagement but still gravitated toward shorter exchanges without the structured session format.

3.3 Experiences that are Valued-but-Avoided

The central observation is a discrepancy between retrospective valuation and prospective choice. Evaluation participants assessed reflective interaction as worthwhile, yet informal deployment showed that without a bounded time commitment and research accountability, users defaulted to the rapid-completion pattern.

The Adaptable dimension illustrates this at finer grain: selected by only 5 of 15 participants, it produced the highest proportion of genuine insight (92% Cognitive Delta). Users did not anticipate the value of contingency thinking. This gap between anticipated and experienced value is itself a product of automation expectation: users optimize for what they expect to need, shaped by the defaults they have learned.

4 Automation Expectation as a Design Problem

If automation expectation systematically forecloses valuable cognitive experiences, it warrants attention in contexts where such foreclosure occurs as a design problem to be addressed rather than a user preference to be accommodated.

4.1 Potential Consequences for Design

Automation expectation operates prior to deliberate choice. The default orientation prevents evaluation of alternatives from occurring, making the status quo self-reinforcing: users cannot prefer what they have not experienced. The downstream consequences are visible in the literature: declining confidence in independent reasoning [9], homogenization of creative outputs [2], and degraded independent capability following sustained AI assistance [7]. Read together with the valued-but-avoided gap we observe, these findings suggest cumulative effects on the capacity for cognitive engagement, particularly in organizational contexts where compliance artifacts have replaced developmental processes [5] and where generative AI accelerates a pattern that predates the technology. Current designs risk affording a perceived cost-benefit that fails to engage users in strategies producing meaningful outcomes. For agentic and composite systems requiring deliberate configuration, this amounts to a substantial adoption barrier.

4.2 Directions for Intervention

Addressing automation expectation requires identifying which contextual sources are most active in a given deployment. We outline directions connected to our use-case, noting that several exceed typical interface design scope.

The limits of reframing the entry-point. Our prototype was entirely question-based and did not offer direct artifact generation. Informal deployment suggests this is insufficient alone: users still attempted to redirect a question-based interface toward artifact production.

Making cognitive outcomes legible. Automation expectation persists partly because the value of reflective engagement is difficult to anticipate. Interfaces that surface cognitive work during interaction, tracking articulations, changed assumptions, or identified constraints, could make the value of sustained engagement visible prospectively.

Designing for premeditative engagement. Users underselected the most insight-generating dimension, suggesting a role for premeditative design [11]: systems that prescribe engagement with dimensions users would not select themselves, calibrated to avoid paternalism while addressing documented biases in self-assessment of cognitive needs [15].

Contextual timing and organizational integration. The structured session succeeded partly through bounded commitment. Integration with calendar-linked interactions or existing reflective practices could replicate the temporal structure that enabled engagement in evaluation but was absent in informal use.

Organizational reinforcement. Interface-level interventions will be limited if organizational environments continue rewarding rapid completion. How deployment policies, time allocation, and evaluation criteria interact with automation expectation requires investigation beyond the interface.

Reframing cognitive effort as productive. In our evaluation, 93% characterized persistent questioning as productive rigor, suggesting the association between cognitive effort and negative experience is not inevitable. If users read “the system is asking me questions” as thinking occurring rather than friction to production, this could begin to shift the expectation. This is distinct from imposing friction; it involves designing interactions where cognitive effort feels generative. It also suggests that automation expectation is as much a cultural and organizational question as an interface design question.

5 Conclusion

We have introduced automation expectation as a term for the convergent default toward rapid delegation that users bring to generative AI. The contributing phenomena are individually documented in both recent human-AI interaction research and the established automation literature; our contribution is in naming their compound effect, distinguishing it from complacency and bias as a pre-engagement orientation, and demonstrating empirically that it forecloses cognitive experiences users value when they occur. The observation derives from a single organizational context in corporate goal-setting, and whether the valued-but-avoided gap generalizes across domains, populations, and task types is an open question. Contexts already structured toward compliance, as ours was, may produce particularly strong automation expectations. We offer this as a conceptual framing and empirical starting point for automation experience research concerned with how automation expectations shape which experiences users will have.

Acknowledgements

We thank Arin Bhowmick, Malini Leveque, and Alex Swain at SAP for providing the research context, contributing to use-case selection, supporting participant access, and their generous engagement throughout the project. This research was funded by SAP in collaboration with DFKI.

References

- [1] Fatemeh Alizadeh, Gunnar Stevens, Oleksandra Vereschak, Gilles Bailly, Baptiste Caramiaux, and Dominik Pins. 2022. Building Appropriate Trust in Human-AI Interactions. In *Proceedings of the 20th European Conference on Computer-Supported Cooperative Work*. https://doi.org/10.48340/ECSCW2022_WS04
- [2] Barrett R. Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization Effects of Large Language Models on Human Creative Ideation. In *Creativity and Cognition*, 413–425. <https://doi.org/10.1145/3635636.3656204>
- [3] Zana Bućinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–21. <https://doi.org/10.1145/3449287>
- [4] Mary L. Cummings, Fei Gao, and Kris M. Thornburg. 2016. Boredom in the Workplace: A New Look at an Old Problem. *Human Factors* 58, 2 (2016), 279–300. <https://doi.org/10.1177/0018720815609503>
- [5] Edward L. Deci, James P. Connell, and Richard M. Ryan. 1989. Self-determination in a Work Organization. *The Journal of Applied Psychology* 74, 4 (1989), 580–590. <https://doi.org/10.1037/0021-9010.74.4.580>
- [6] Cedric Faas, Richard Bergs, Sarah Sterz, Anna Maria Feit, and Markus Langer. 2025. Give Me a Choice: The Consequences of Restricting Choices Through AI-Support for Perceived Autonomy, Work Meaningfulness, and Decision Performance. In *Proceedings of the AutomationXP25 Workshop at CHI'25*.
- [7] Andreas Fügener, Jörn Grahl, Alok Gupta, and Wolfgang Ketter. 2022. Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research* 33, 2 (2022), 678–696. <https://doi.org/10.1287/isre.2021.1079>
- [8] Gary Klein, David D. Woods, Jeffrey M. Bradshaw, Robert R. Hoffman, and Paul J. Feltovich. 2004. Ten Challenges for Making Automation a “Team Player” in Joint Human-Agent Activity. *IEEE Intelligent Systems* 19, 6 (2004), 91–95. <https://doi.org/10.1109/MIS.2004.74>
- [9] Hao-Ping (Hank) Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. 2025. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects from a Survey of Knowledge Workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–22. <https://doi.org/10.1145/3706598.3713778>
- [10] Raja Parasuraman and Dietrich H. Manzey. 2010. Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors* 52, 3 (2010), 381–410. <https://doi.org/10.1177/0018720810376055>
- [11] Ronda Ringfort-Felner, Judith Dörrenbächer, and Marc Hassenzahl. 2025. The Quality of Speculation – A Scoping Review. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, 2373–2394. <https://doi.org/10.1145/3715336.3735794>
- [12] Evan F. Risko and Sam J. Gilbert. 2016. Cognitive Offloading. *Trends in Cognitive Sciences* 20, 9 (2016), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
- [13] Tobias Sturm, Luisa Pumplun, Matthias Relke, and Peter Buxmann. 2025. Under Pressure: Contextualizing Workplace Stress Towards User-Centered Interventions. *Behaviour & Information Technology* (2025). <https://doi.org/10.1080/0144929X.2025.2587732>
- [14] Hari Subramonyam, Roy Pea, Christopher Pondoc, Maneesh Agrawala, and Colleen Seifert. 2024. Bridging the Gulf of Envisioning: Cognitive Challenges in Prompt Based Interactions with LLMs. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3613904.3642754>
- [15] J. Lukas Thürmer and Frank Wieber. 2015. The Planning Fallacy: When Plans Lead to Optimistic Forecasts. In *The Psychology of Planning in Organizations*. Routledge, 47–73. <https://doi.org/10.4324/9780203105894-7>
- [16] Dhaval Vyas, Cristina M. Chisalita, and Alan Dix. 2016. Organizational Affordances: A Structuration Theory Approach to Affordances. *Interacting With Computers* 29, 2 (2016), 117–131. <https://doi.org/10.1093/iwc/iww008>
- [17] Justin D. Weisz, Jessica He, Michael Muller, Gabriela Hoefler, Rachel Miles, and Werner Geyer. 2024. Design Principles for Generative AI Applications. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–22. <https://doi.org/10.1145/3613904.3642466>
- [18] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376301>
- [19] J. D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny Can’t Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–21. <https://doi.org/10.1145/3544548.3581388>
- [20] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research Through Design as a Method for Interaction Design Research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/1240624.1240704>