Agents act, users don't

A reflection on possible hidden costs of agentic technology from a psychological perspective

Lara Christoforakos*

Department Psychology, LMU Munich, lara.christoforakos@psy.lmu.de

Sarah Diefenbach

Department Psychology, LMU Munich, sarah.diefenbach@psy.lmu.de

Automation is increasingly taking over everyday tasks. With technological development automation is becoming ever more agentic, meaning that systems act in a proactive, independent and self-governed manner. While such developments promise benefits in convenience, performance and overall efficiency, we argue that they introduce underexplored psychological risks on an individual level. First, agentic autonomous systems may evoke psychological reactance by threatening perceived behavioral freedom, thereby reducing user acceptance and compliance with the system. Second, increasing agency in automation may shift responsibility attribution away from users, diminishing perceived personal accountability for outcomes of HCI. Third, agentic automation may foster borrowed competence, whereby users experience reduced skill development alongside erroneous self-perceptions of competence, increasing dependence on technology over time. It is therefore essential for HCI research to explore how automation can be designed for users to optimally derive advantages while simultaneously ensuring technology acceptance, appropriate responsibility attribution for outcomes and user skill development.

CCS CONCEPTS • Human-centered computing • Human computer interaction (HCI) • HCI theory, concepts and models

Additional Keywords and Phrases: Agentic automation, psychological reactance, responsibility attribution, skill loss

ACM Reference Format:

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, [Number] pages.

1 INTRODUCTION

Nowadays automation is increasingly spreading into various aspects of everyday life [8] . From robotic vacuum cleaners and smart thermostats in the household, to automated cars and public transport, technological systems are ever more assuming responsibility for tasks that previously required not only manual human effort but also ongoing human oversight. It is not to be denied that this comes with obvious advantages. A central benefit is that it liberates time for activities and interactions of great individual significance [6]. Moreover, the wide implementation of automation generally improves efficiency and precision, while lowering costs. In addition, vacuum cleaners, smart thermostats or self-driving cars do not simply execute user commands, they can also proactively decide, when the floor is dusty enough to start cleaning, which

* Place the footnote text for the author (if applicable) here.

temperature is appropriate for an environment-friendly heating or which route to drive, depending on the current traffic situation, just to name a few mundane examples. Thus, with increasing automation, technologies are also becoming more agentic, meaning that they are proactive, independent, and self-governed [17]. While a high agency in automation can generally further support the effectiveness of such technological concepts, it can also come with hidden costs on a societal as well as individual level. On a societal level such hidden costs have been often discussed, mainly considering the question of whether automation will replace jobs performed by humans [19]. The more subtle psychological costs, which are mainly observable on an individual level in everyday life, are often overlooked. In this regard, we see at least three potential risks of highly agentic automation.

First, although highly agentic automation is intended to relieve users of routine or undesirable tasks (e.g., vacuuming), systems that independently initiate actions, such as an agentic shower system which automatically lowers water temperature after a certain amount of time to support environment-friendly behavior, may elicit **psychological reactance**. Psychological reactance is an aversive motivational state arising from a perceived threat to, or loss of, one's behavioral freedom [2]. Consequently, users may exhibit reduced acceptance of the system or engage in behaviors that oppose the technology's recommendations [13].

A further potential risk of highly agentic automation concerns **shifts in responsibility attribution** (e.g., [4, 6]). As technological agency increases, users may attribute less responsibility to themselves for the outcomes of human–technology interaction. For example, when a surgical robot autonomously determines which surgical procedure is the best fit for a patient, human surgeons may feel less responsible for the outcome of that decision, even though they selected, configured, and authorized the system's use. If the intervention later harms the patient or fails to deliver the expected benefit, surgeons may also be less likely to perceive themselves as accountable for this negative outcome.

Finally, when agentic automation assists users in achieving specific goals, it may inadvertently **impede the development of users' goal-directed** strategies (e.g.,[12]), thereby fostering dependence on the continued presence of the technology. For example, individuals who rely on a smart thermostat to reduce household energy consumption may fail to acquire or practice environment-friendly behaviors, as the automation performs the relevant regulatory functions on their behalf. As a result, users may be unable to show the corresponding behaviors in the absence of the technology, potentially also leading to diminished self-efficacy. While issues of responsibility attribution and competence development may apply to artificial intelligence systems more broadly, they become particularly salient in the context of agentic automated systems, as these systems increasingly perform tasks on behalf of users. Additionally such systems may enforce reactant user behavior as they come with the potential to restrict user control and behavioral freedom. In the following we elaborate on the above-named phenomena and related previous research in more detail.

2 THE ISSUE OF PSYCHOLOGICAL REACTANCE

When automation operates in a highly agentic manner, characterized by increased proactivity (initiating actions without user input), self-governance and independence (making decisions independently and modifying behavior based on inferred user states) it may alter how users perceive and respond to the system. In such cases, automation does not merely execute predefined commands but actively intervenes in users' environments or behavioral routines. These agentic characteristics can evoke psychological reactance [2], leading users to resist or even oppose the system. Psychological reactance, originally defined by Brehm [2] as an unpleasant motivational arousal arising from perceived threats to behavioral freedom, functions to restore that threatened freedom [2]. Accordingly, the more agentic automation becomes the more it may heighten perceived threats to behavioral freedom and thereby intensify reactance. In human-computer interaction (HCI) literature, Roubroeks et al. [18] also argue, drawing on social agency theory, that social cues in technological systems can activate social schemata, suggesting that reactance may be stronger when a freedom threat is attributed to a social agent. Automation characterized by high proactivity, self-governance, and independence may therefore be perceived as more agent-like and elicit stronger reactance than less agentic systems.

Empirical findings support this assumption. Jensen et al. [9] report that users interacting with automated electricity-shifting systems expressed concerns about losing control due to system automation. Similarly, in studies of automated heating systems, participants reported irritation with system behavior and often constrained their temperature preferences to limit automation [10], indicating attempts to counter perceived autonomy loss. Furthermore, in a recent study [3] in which a smart-home system either suggested or autonomously implemented sustainable behaviors, higher technology agency (automated adjustments vs. suggestions) reduced both pro-environmental behavioral intentions and intentions to use the system, with psychological reactance mediating these effects. Together, these findings suggest that psychological reactance represents a critical risk factor for agentic technologies designed to support users in everyday tasks.

3 THE ISSUE OF RESPONSIBILITY ATTRIBUTION

Weiner's Theory of Social Conduct explains how attributions of agency and responsibility emerge in interpersonal contexts, conceptualizing agency as an antecedent of perceived responsibility [21]. Thereby, agency describes the capacity to act independently and implies controllability. Responsibility, in turn, requires both a controllable cause and intentional action [21]. Transferred to the context of agentic automation, these mechanisms become particularly relevant as systems that exhibit proactivity, self-governance and independence, for example in decision-making, may increasingly be perceived as actors with intentionality and control. As a result, higher levels of system agency can elevate the degree to which responsibility is attributed to the system itself. Empirical work suggests that these mechanisms extend to human-technology interaction. Van Der Woerdt and Haselager [22], for example, found that individuals attributed greater responsibility to robots perceived as more agentic, a pattern frequently observed across HCI research on autonomous and anthropomorphic systems [5, 11, 22, 23].

At the same time, increased technology agency may shift responsibility away from users. Limerick et al. [14] argue that as system autonomy increases and the user's role becomes less visible, users' self-attributed responsibility may decline. Empirical findings support this argument. In a recent study within the household context, users reported lower responsibility for outcomes when tasks were performed by robotic rather than manual cleaning systems [6]. Comparable effects have been observed in other domains, such as automated parking assistance [15]. Similarly, results of vignette-based study indicate that highly agentic smart-home systems that autonomously adjusted temperature settings reduced users' perceived responsibility for resulting energy savings compared to systems that merely provided suggestions, which in turn diminished users' intention to align with the system's goals, i.e., chose environment-friendly settings within the smart-home system [4]. In sum, shifts in responsibility might be a central issue with agentic automation, possibly coming with behavioral consequences.

4 THE ISSUE OF BORROWED COMPETENCE

When automation performs tasks on behalf of users, it may reduce opportunities for skill acquisition while simultaneously fostering an inflated perception of one's own competence. This dynamic is particularly pronounced in the context of agentic automation, where systems not only assist but may proactively and independently substitute user action. By taking over task execution, such systems can create an illusory sense of mastery while diminishing actual competence due to reduced opportunities for practice and feedback [1]. As a result, users may become increasingly dependent on the technology and experience performance decrements when the system is unavailable. Within HCI literature, such processes have been linked to deskilling [1, 16] and overreliance [20], suggesting that perceived user competence may not equal actual user competence, when highly agentic automation is involved.

Empirical research in this regard has, for example, explored neural and behavioral consequences of LLM-assisted essay writing. Findings show that greater reliance on external tools, especially LLMs, was associated with lower brain connectivity, reduced cognitive engagement, weaker memory for one's own writing, and lower perceived ownership of essays, with participants who did not use external tools showing the strongest neural and behavioral performance. When tool conditions were switched, prior LLM users remained under-engaged while participants who had not used external tools before showed stronger activation patterns, suggesting potential cognitive costs from sustained LLM reliance [12]. Consequently, borrowed competence represents a potential risk of agentic automation, as short-term gains in user performance may come with the cost of longer-term erosion of users' skills.

5 OUTLOOK

While highly agentic automation offers substantial individual and societal benefits, it may also come with psychological challenges. Focusing on the individual level, support from highly agentic systems may elicit opposing behaviors when users perceive threats to their behavioral freedom, reduce users' perceived responsibility for outcomes achieved with technological assistance, and foster illusory competence when performance primarily roots in the system rather than the user. At the individual level, these risks may diminish willingness to engage with agentic systems and undermine self-efficacy and motivation [7], with potential consequences for overall well-being. At the societal level, such dynamics may complicate the adoption of automation in domains related to everyday life, including household and professional contexts. Shifts in responsibility attribution and miscalibration of users' perceived skills may contribute to broader skill erosion and an increasing human dependence on technological support. Moreover, it may create an unsustainable demand for the continual anticipation and monitoring of behaviors and tasks that could require technological support, alongside the ongoing development of corresponding systems. Accordingly, key questions for responsible HCI design arise. A first question concerns how to calibrate automation agency to support users in daily life without provoking reactance or undermining user capability. Addressing this question requires understanding thresholds of acceptable agency in automation, as well as how system design features can affect perceived behavioral freedom of users, closely linked to psychological reactance and the perception of own capability. Second, how can agentic systems be designed to preserve or even strengthen users' sense of responsibility? This raises questions about how to make user contributions visible despite autonomous systems, how to design for shared agency between human and system, and how responsibility for outcomes can be optimally communicated in interaction. Finally, another key question concerns how agentic systems might even foster the development of long-lasting user competences across domains. This includes exploring when systems should intervene versus defer to the user, how to incorporate handover of control to the user, and how feedback mechanisms can maintain users' learning and accurate self-assessment over time.

ACKNOWLEDGMENTS

Parts of this research have been funded by the Ministry of Research, Technology and Space (Research Project: MOVEN; 01UU2204B).

REFERENCES

[1] Matthias Baldauf, Peter Fröehlich, and Rainer Endl. 2020. Trust Me, I'm a Doctor – User Perceptions of AI-Driven Apps for Mobile Health Diagnosis. In Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia (MUM '20). Association for Computing Machinery, New York, NY, USA, 167–178. https://doi.org/10.1145/3428361.3428362

[2] Jack W. Brehm. 1966. A theory of psychological reactance. Academic Press.

[3] Lara Christoforakos, Till M. Jürgens, Melina Busch, Leonie Beuckelmann, Matthias Laschke, and Sarah Diefenbach. 2026. When Technology's Well-Meant Recommendations Backfire. Exploring the Role of Agency on Psychological Reactance in Technology for Sustainable Behavior Change. Behaviour &Information Technology, in press. https://doi.org/10.1080/0144929X.2026.2628766.

[4] Lara Christoforakos, Alina Kolb, and Sarah Diefenbach. 2024. I did it once, but can I do it again? The role of responsibility attribution and self-efficacy in technology for sustainable behavior change. In Proceedings of the 13th Nordic Conference on Human-Computer Interaction (NordiCHI '24). Association for Computing Machinery, New York, NY, USA, Article 6, 1–15. https://doi.org/10.1145/3679318.3685339

[5] Lara Christoforakos, Ewgenia Borodkow, and Sarah Diefenbach. 2020. Who gets the fame, who is to blame? Empirical exploration of responsibility attribution in HCI. In T. Köhler, E. Schoop, & N. Kahnwald (Eds.), Gemeinschaften in Neuen Medien. Von hybriden Realitäten zu hybriden Gemeinschaften. Communities in New Media. From hybrid realities to hybrid communities. 226–235. TUDpress.

[6] Sarah Diefenbach, Daniel Ullrich, Tim Lindermayer and Kaja-Lena Isaksen. 2024. Responsible Automation: Exploring Potentials and Losses through Automation in Human–Computer Interaction from a Psychological Perspective. Information 15, 8 (August 2024), 460. https://doi.org/10.3390/info15080460

[7] Patrick Ehrenbrink, Xin Guang Gong, and Sebastian Möller. 2016. Implications of different feedback types on error perception and psychological reactance. In Proceedings of the 28th Australian Conference on Computer-Human Interaction (OzCHI '16). Association for Computing Machinery, New York, NY, USA, 358–362. https://doi.org/10.1145/3010915.3010994

[8] Christian P. Janssen, Stella F. Donker, Duncan P. Brumby, and Andrew L. Kun. 2019. History and future of human-automation interaction. International Journal of Human Computer Studies 131 (November 2019), 99–107. https://doi.org/10.1016/j.ijhcs.2019.05.006

[9] Rikke Hagensby Jensen, Jesper Kjeldskov, and Mikael B. Skov. 2018. Assisted Shifting of Electricity Use: A Long-Term Study of Managing Residential Heating. ACM Trans. Comput.-Hum. Interact. 25, 5 (October 2018), 1–33. https://doi.org/10.1145/3210310

[10] Rikke Hagensby Jensen, Jesper Kjeldskov, and Mikael B. Skov. 2016. HeatDial: Beyond User Scheduling in Eco-Interaction. In Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16). Association for Computing Machinery, New York, NY, USA, Article 74, 1–10. https://doi.org/10.1145/2971485.2971525

[11] Taemie Kim and Pamela Hinds. 2006. Who Should I Blame? Effects of Autonomy and Transparency on Attributions in Human-Robot Interaction. In ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication, 80–85. https://doi.org/10.1109/ROMAN.2006.314398

[12] Nataliya Kosmyna, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes. 2025. Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv preprint arXiv:2506.08872, 4.

[13] Matthias Laschke, Sarah Diefenbach, and Marc Hassenzahl. 2015. "Annoying, but in a nice way": an inquiry into the experience of frictional feedback. International Journal of Design 9, 2 (August 2015), 129-140.

[14] Hannah Limerick, David Coyle, and James W. Moore. 2014. The experience of agency in human-computer interactions: A review. Frontiers in Human Neuroscience 8, 643 (August 2014), 1–10. https://doi.org/10.3389/fnhum.2014.00643

[15] Elisabeth Moosheimer and Sarah Diefenbach. 2026. Responsible Automation. Exploring the impact of device automation on responsibility for the result in a field study [Manuscript submitted for publication]. Department of Psychology, Ludwig-Maximilians Universität München.

[16] Chiara Natali, Luca Marconi, Leslye D. Dias Duran and Federico Cabitza. 2025. AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond. Artificial Intelligence Review 58, 356 (August 2025). https://doi.org/10.1007/s10462-025-11352-1

[17] Thomas O'Neill, Nathan McNeese, Amy Barron, and Beau Schelble. 2022. Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. Human Factors: The Journal of the Human Factors and Ergonomics 64, 5 (August 2022), 904–938. https://doi.org/10.1177/0018720820960865

[18] Maaike Roubroeks, Jaap Ham, and Cees Midden. 2011. When artificial social agents try to persuade people: The role of social agency on the occurrence of psychological reactance. International Journal of Social Robotics 3, 2 (January 2011), 155–165. https://doi.org/10.1007/s12369-010-0088-1

[19] Alana Semuels. 2020. Millions of Americans Have Lost Jobs in the Pandemic—And Robots and AI Are Replacing Them Faster Than Ever. Time (6 August 2020). Retrieved April 11, 2024 from https://time.com/5876604/machines-jobs-coronavirus/

[20] Prakash Shukla, Phuong Bui, Sean S Levy, Max Kowalski, Ali Baigelenov, and Paul Parsons. 2025. De-skilling, Cognitive Offloading, and Misplaced Responsibilities: Potential Ironies of AI-Assisted Design. In Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25). Association for Computing Machinery, New York, NY, USA, Article 171, 1–7. https://doi.org/10.1145/3706599.3719931

[21] Bernard Weiner. 1995. Judgments of Responsibility: A Foundation for a Theory of Social Conduct. Guilford Press.

[22] Sophie Van Der Woerdt and Pimm Haselager. 2019. When robots appear to have a mind: The human perception of machine agency and responsibility. New Ideas in Psychology 54 (August 2019), 93–100. https://doi.org/10.1016/j.newideapsych.2017.11.001

[23] Adam Waytz, Joy Heafner, and Nicholas Epley. 2014. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. Journal of Experimental Social Psychology 52, (May 2014), 113–117. https://doi.org/10.1016/j.jesp.2014.01.005