

Do Generative AI Tools Foster Positive Experiences in Knowledge Work? A NeuroIS Research Proposal



Michael Thomas Knierim and Lorenzo del Puppo

Abstract Generative AI has seen a significant rise in its performance and use in various fields, including academia, art, design, and software engineering. However, little research has been conducted on how users interact with AI tools. This article proposes a research project that focuses on the elicitation of positive psychological experiences, such as flow, when an individual is fully immersed in a task and feels highly effective and satisfied. The study aims to investigate how AI tools may promote flow experiences by regulating cognitive load to flow-conducive levels through providing initial solutions to demanding tasks or shifting user demands from monotonous tasks to more challenging ones. The study will utilize wearable EEG recordings to generate objective insights into the dynamics of cognitive load and flow during early project stages. The research could lead to the development of neuro-adaptive recommender agents that propose AI invocation when undesirable load levels are detected.

Keywords Generative AI · ChatGPT · Cognitive load · Flow · Wearable EEG

1 Introduction

Artificial Intelligence has been on the rise for several years [1, 2]. Yet, with the recent emergence of generalized transformer models like ChatGPT a significant shift in their performance has occurred [3]. In academia, prestigious journals like Nature and Science have already highlighted the rise of Human-AI-collaborative projects [3, 4]. Also, other powerful generative tools for image (e.g. DALL-E), audio (e.g. Speechify), and video generation (e.g. Synthesia), but also coding and software application development (e.g. Uizard) have been developed and impact how people

M. T. Knierim (✉) · L. del Puppo
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
e-mail: michael.knierim@kit.edu

L. del Puppo
e-mail: lorenzo.puppo@kit.edu

produce innovative media and knowledge. While some research has focused on the technical capabilities and limitations [5, 6] and socio-economic impacts on the workforce [1, 3], little attention has been paid to the user experience of working with these powerful AI tools. However, as the popularity of ChatGPT has demonstrated (it gained over 100 million users in less than two months), it is crucial to examine how users interact with these tools and the experiences they elicit.

Anticipating a future where people interact with AI tools rather than being replaced by them, it is pertinent to explore how it feels to work with them [7]. This study proposes to focus on one possible user experience outcome: the elicitation of positive psychological experiences [8]. Specifically, we aim to investigate how AI tools may promote experiences of mastery and well-being by facilitating flow, a state of optimal experience characterized by a high level of task engagement and satisfaction [9]. This flow experience has been an emerging topic in IS—and particularly in NeuroIS research (Koufaris, 2002; Léger et al., 2014; Nadj et al., 2023; Knierim et al., 2017; Knierim, 2018; [10–16]). Since flow occurs when skills meet task demands (and when both are present at high levels) [17], and since cognitive load has been identified as a proxy for this skill-demand balance [18], we propose that generative AI tools could regulate cognitive load to flow-conducive levels, for example, by providing initial solutions to challenging tasks (lowering task demands) or by shifting user demands from monotonous tasks to more challenging ones (increasing task demands).

In this article, we propose a study design using NeuroIS methodology to investigate these patterns in a knowledge work task scenario. By using wearable EEG recordings, we also aim to gain objective insights into the dynamics of cognitive load and flow, especially during early project phases. As the EEG is well known for monitoring correlates of cognitive load [19] with high temporal resolution, this approach also allows for theoretical extensions of load-flow dynamics in real-time. In the following sections, we detail the theoretical foundations of flow and cognitive load that we use to develop a corresponding study design. Ultimately, this research could lead to neuro-adaptive recommendation agents that suggest AI invocation when undesirable load levels are detected—and ideally facilitate positive experiences when doing so. Before this happens, however, additional questions need to be answered, which are discussed in the outlook section of this article.

2 Related Work

Flow and Related IS Work. The concept of flow encompasses nine dimensions that include the perception of: (1) a balance between task demand and an individual’s skill, (2) clear goals, (3) unambiguous feedback, as well as (4) effortless concentration, (5) merging of action and awareness, (6) loss of self-consciousness, (7) control, (8) distortion of time, and (9) intrinsic reward [9]. Research has shown that flow can be experienced in any task that requires active engagement and fulfills the three pre-conditions, of which the balance between perceived demand and skill is particularly noteworthy, as it is often manipulated in flow experiments [17]. The concept of flow

has been extensively studied in business and IS research, with several studies linking it to improved job performance, creativity, and enjoyment [20–22]. Additionally, it has been associated with technology-enabled team building [23] as well as technology adoption and use [11, 16], highlighting its potential impact on desirable outcomes in the context of IS use and work. The relevance of flow experiences in IS use scenarios and knowledge work has also been recognized in NeuroIS research [10]. While much of the previous research has focused on identifying neurophysiological correlates of flow experiences [14, 15], recent studies have highlighted the connection between flow and cognitive load, which serves as the foundation for this research proposal.

Flow and Cognitive Load. In general, cognitive processing is composed of two primary components: a relatively limited working memory and a much larger long-term memory [24]. When engaging in a task, some degree of working memory must be allocated to that task, a process referred to as cognitive load [24, 25]. Previous work has found that cognitive load and flow experiences are linked through an inverted U-shaped relationship [18, 26]. This has been documented both for self-reports and for well-known EEG correlates of cognitive load like frontal Theta, and posterior Alpha and Beta frequency band power changes [19]. The explanation for this relationship is primarily rooted in the flow pre-condition of demand-skill balance [18]. When a task is neither too demanding nor too easy, the efficient and automated task processing that characterizes flow can occur, likely because processes that are detrimental to the primary task are downregulated (e.g. self-referential attention, conflict monitoring or mind-wandering—[27]). Beyond this general integration of the two theories, little is reported about their cognitive and temporal dynamics. However, this represents an interesting aspect because (1) scholars have reported that flow often emerges sporadically and chaotically [28], and (2) an understanding of the temporal dynamics would provide a highly valuable foundation for the development of flow-facilitating technology (e.g. by understanding when and how to invoke feedback—see [12]). Therefore, in investigating the impact of human-AI collaboration on positive experiences, we propose using NeuroIS methods to study these temporal dynamics to enable theoretical contributions beyond observing their outcomes.

3 Study Proposal

Study Concept, Hypotheses and Procedure. Our proposed study investigates the impact of generative AI tools, such as ChatGPT, on positive experiences in a knowledge work setting. To achieve this, we build on previous NeuroIS research that has examined the emergence of flow experiences in knowledge work tasks like scientific writing [15, 29]. One notable finding from this work was that cognitive load peaked during the start of the writing stage and then decreased after a few minutes [29]. This aligns with flow and writing research that suggests the start of a writing session may require more effort to structure the task than later stages [9, 30]. Based on this observation, we propose that the early stages of a writing project are particularly

demanding and that this moment presents an opportunity to use an AI tool to reduce task demands. Thus, we derive the following hypotheses:

- H1:** Cognitive load increases during early stages of a complex writing task.
- H2:** Using generative AI reduces cognitive load during early stages of a complex writing task.
- H3:** The reduction of cognitive load in an early writing stage increases the intensity of flow experiences.

Furthermore, because it is known in flow literature that task skill and importance (here: topic knowledge) significantly influence when and how flow is experienced [31], we expect these two factors to moderate hypotheses H2-H3:

- H4a:** The more important a topic is to a writer, the greater the AI support’s load reduction and flow intensification.
- H4b:** The more knowledge a writer has with a topic, the weaker the AI support’s load reduction and flow intensification.

To account for individual differences in EEG data and skill levels, we propose a fully within-subject experimental design that includes (1) a standardized text copying task of varying difficulty, and (2) drafting two extended abstracts on a standardized topic, once with and once without the support of a generative AI tool. Figure 1 illustrates the proposed experimental structure, including the procedure, tasks, and questionnaires.

Tasks and Manipulations. The text copying task (see, e.g. [32]) is proposed as a baseline measure of cognitive load to resemble common knowledge work. This text copying task consists of a set of pre-selected text segments of varied lengths that are presented to participants for 10 s per segment. By varying the length of each segment

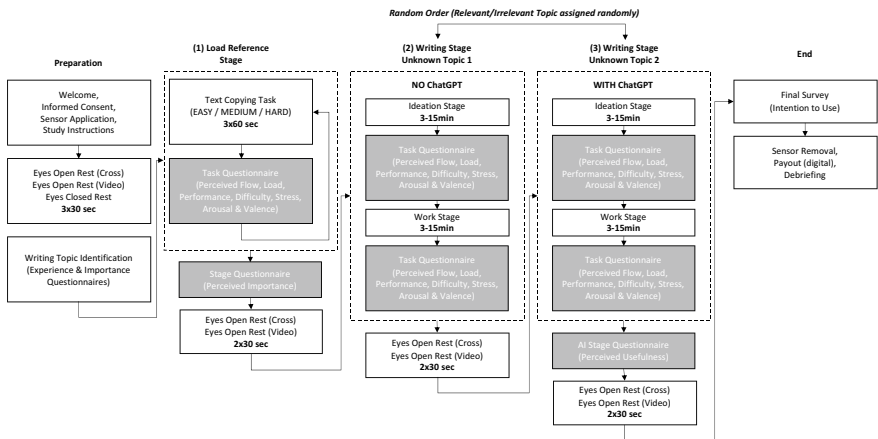


Fig. 1 Procedure of the proposed experiment

that has to be typed, low, medium, and high levels of cognitive load can be elicited that can serve as a reference for the report and EEG data later (see, e.g. [18]).

For the abstract writing stage, participants will work on two standardized topics that are identified in an early stage of the experiment. Specifically, as we expect topic relevance and expertise to moderate the load and flow changes induced by generative AI tools (H4a and H4b), participants are asked to indicate this relevance and expertise for a pre-defined range of four topics at the start of the experiment. These four topics will cover recent developments in the IS discipline. To maximize contrasts, participants will be assigned to work on the topic with the highest and the lowest relevance and expertise—randomly assigned to the AI support/no support conditions. Further, to compare general ability levels for a writing task, we will include the writing practice index (WPI) used in related work on creative writing [33]. The order of the two writing assignments will be randomized to counter any unwanted carryover effects.

The writing process itself is divided into two phases. First, in the ideation stage, participants will have time to draft a storyline for their abstract without the support of any AI tool. They will be asked to focus on generating a story in bullet points. We consider this phase necessary because it will serve to test the hypotheses that cognitive load increases substantially in the early stages of a complex writing task (H1), and that the use of a supportive generative AI tool thereafter reduces cognitive load and flow experience (H2 and H3). In addition, the use of the generative AI tool is more comparable across participants because it is not used for ideation, but rather for implementation of the writing. Second, in the work phase, participants are asked to complete the writing of their text based on their previous ideas. In one round, they are asked to use ChatGPT to produce their extended abstract. In both phases, participants must work for a minimum of three minutes and a maximum of 15 min. We offer this variable time span to account for individual differences in the time it takes participants to produce a satisfactory solution. After each ideation and work stage, participants are asked to fill out surveys about their load and flow experience levels—similar to previous work [15, 34, 35].

To complete the experiment, participants will complete resting baseline measurements for 30 s with eyes open (fixating on a cross) and 30 s watching a video of fish swimming in the ocean as a more natural resting stage stimulus [36]. Overall, we expect the experiment to last 60 min, including the time for instruction and set-up of the instrumentation. Participants will be remunerated with a flat fee following recommendations for flow research [37].

Measures. The NASA Task Load Index (six items by [38]) and the short flow scale (ten items by [31]) will be used as main reports at each task interruption or conclusion. In addition, at each interruption, we will include additional items to assess the task difficulty (one item by [31]) and other facets of the affective experience during the task (a five-item stress measure by [39] and two pictorial items for emotional arousal and valence—[40]). After each task (the three main stages), we will also ask about the task importance (three items by [31]) to compare if task motivation might have influenced the experiences. Finally, at the end of the experiment, participants will



Fig. 2 Dry-electrode wearable EEG system for cognitive load monitoring (see [18])

be asked about their intention to use the AI in similar situations again (three-item TAM construct by [41]—for the AI use condition), or whether they would like to use a generative AI for such tasks (for the non-AI use condition—after description of ChatGPT and its capabilities—adapting the three TAM construct items again).

To continuously monitor cognitive load levels, we will use a wearable EEG system with dry electrodes on the top of the head (importantly with electrodes at Cz)—specifically, a system that resembles headphones that could be used in everyday life (see, e.g. [18, 42] and Fig. 2). We opt for such a system because we believe it is important that NeuroIS scholars engage with such everyday life systems to advance the development of adaptive systems. As cognitive load effects are large and well-observable over frontal and central midline positions [19] we also believe such a system to be an appropriate choice in terms of rigor. We expect to see the classic increase in Theta frequency band power and decrease in Alpha frequency band power as cognitive load increases, especially over the C3, Cz, and C4 positions.

4 Discussion and Outlook

In our work, we propose a study that aims to investigate the impact of generative AI tools on cognitive load and flow experience. We believe that understanding the effects of emerging technologies on user experiences is of paramount importance, especially since the integration of AI tools has the potential to enhance productivity and enjoyment in the future of work through human-AI collaboration. Our proposed study design incorporates wearable EEG systems to observe cognitive dynamics with high temporal resolution, which will contribute significantly to the theoretical integration and extension of load and flow theory. We hope that our findings will provide a novel foundation for explaining how cognitive load patterns contribute to the emergence of flow experiences and how they are regulated by AI tools. It is important to note that our approach is based on the assumption that AI tools provide useful solutions that can be easily integrated into primary tasks. However, if this assumption does not hold, the use of AI tools may have the opposite effect,

increasing load and creating more stressful experiences. This possibility highlights the importance of studying the human-AI interaction in knowledge work scenarios, which are anticipated to occur at increasing rates now that generalized transformer models have made their debut in the modern world. As IS scholars, we have an opportunity here to lead the way in studying and supporting the impact of AI on people, organizations, and society as a whole.

References

1. Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*, 254–280.
2. Perotti, J. W., Wall, P., & McLaughlin, G. (2010). The future of knowledge work: Predictions for 2020. *Horiz*, *18*, 213–221. <https://doi.org/10.1108/10748121011072663>
3. Thorp, H. H. (2023). ChatGPT is fun, but not an author
4. Editorial, N. (2023). Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature*, *613*, 612.
5. Dowling, M., & Lucey, B. (2023). ChatGPT for (Finance) research: The Bananarama conjecture. *Finance Research Letters*, 103662. <https://doi.org/10.1016/j.frl.2023.103662>.
6. Rudolph, J., Tan, S., & Tan, S.: ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, *6*, 23. <https://doi.org/10.37074/jalt.2023.6.1.9>.
7. Hemmer, P., Westphal, M., Schemmer, M., Vetter, S., Vössing, M., & Satzger, G. (2023). Human-AI collaboration: The effect of AI delegation on human task performance and task satisfaction. In: *IUI '23: Proceedings of the 28th International Conference on Intelligent User Interfaces* (pp. 453–463). <https://doi.org/10.1145/3581641.3584052>.
8. Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, *55*, 5–14. <https://doi.org/10.1037//0003-066X.55.1.5>
9. Csikszentmihalyi, M. (1996). *Flow and the psychology of discovery and invention*. Harper Collins.
10. Riedl, R., Fischer, T., Léger, P.-M., & Davis, F. D. (2020). A decade of NeuroIS research: Progress, challenges, and future directions. In: *ACM SIGMIS Database DATABASE Adv. Inf. Syst.* (Vol. 51, pp. 13–54).
11. Koufaris, M. (2002). Applying the technology acceptance model and flow theory to online consumer behavior. *Information Systems Research*, *13*, 205–223.
12. Nadj, M., Rissler, R., Adam, M.T.P., Knierim, M.T., Li, M.X., Mädche, A., Riedl, R.: What disrupts flow in office work? A NeuroIS study on the impact of frequency and relevance of IT-mediated interruptions. *MIS Q.* in press (2023).
13. Léger, P. M., Davis, F. D., Cronan, T. P., & Perret, J. (2014). Neurophysiological correlates of cognitive absorption in an enactive training context. *Computers in Human Behavior*, *34*, 273–283. <https://doi.org/10.1016/j.chb.2014.02.011>
14. Knierim, M. T., Rissler, R., Dorner, V., Maedche, A., & Weinhardt, C. (2017). The psychophysiology of flow: A systematic review of peripheral nervous system features. In: *Proceedings of the 9th Retreat on NeuroIS* (pp. 109–120).
15. Knierim, M. T., Rissler, R., Hariharan, A., Nadj, M., & Weinhardt, C. (2018). Exploring flow psychophysiology in knowledge work. In: *Proceedings of the 10th Retreat on NeuroIS* (pp. 239–249).
16. Rissler, R., Nadj, M., & Adam, M. T. P. (2017). Flow in Information Systems research: Review, integrative theoretical framework, and future directions. In: *International Conference on Wirtschaftsinformatik* (pp. 1051–1065).

17. Moneta, G. B. (2021). On the conceptualization and measurement of flow. In C. Peifer & S. Engeser (Eds.), *Advances in flow research* (pp. 31–69). Springer Cham. https://doi.org/10.1007/978-3-030-53468-4_2.
18. Knierim, M. T., Zimny, C., & Seitz, J. (2022). Monitoring cognitive load with 3d-printed EEG headphones. In: *43rd International Conference on Information Systems (ICIS)* (p. 17).
19. Chikhi, S., Matton, N., & Blanchet, S. (2022). EEG power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology*, 1–24.
20. Bakker, A. B., & van Woerkom, M. (2017). Flow at work: A self-determination perspective. *Occupational Health Science*, 1, 47–65.
21. Knierim, M. T., Nadj, M., Li, M. X., & Weinhardt, C. (2019). Flow in knowledge work groups—Autonomy as a driver or digitally mediated communication as a limiting factor? In: *Proceedings of the 40th International Conference on Information Systems (ICIS)* (pp. 1–17).
22. Spurlin, S., & Csikszentmihalyi, M. (2017). Will work ever be fun again? In C. J. Fullagar & A. Delle Fave (Eds.), *Flow at work: Measurement and implications* (pp. 176–187).
23. Keith, M., Anderson, G., Dean, D., Gaskin, J., Wisner, T., Gremmert, S., & Payne, K. (2016). The effects of video gaming on work group performance. In: *Proceedings of the 37th International Conference on Information Systems (ICIS)* (pp. 1–20).
24. Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychology*, 38, 1–4.
25. Grimes, M. G., & Valacich J. S. (2015). Mind over mouse: The effect of cognitive load on mouse movement behavior. In: *Proceedings of the 36th International Conference on Information Systems (ICIS)* (pp. 1–13).
26. Ewing, K. C., Fairclough, S. H., & Gilleade, K. (2016). Evaluation of an adaptive game that uses EEG measures validated during the design process as inputs to a biocybernetic loop. *Frontiers in Human Neuroscience*, 10, 1–13.
27. Harris, D. J., Vine, S. J., & Wilson, M. R. (2017). Neurocognitive mechanisms of the flow state. *Progress in Brain Research*, 237, 221–243. <https://doi.org/10.1016/bs.pbr.2017.06.012>
28. Ceja, L., & Navarro, J. (2012). ‘Suddenly I get into the zone’: Examining discontinuities and nonlinear changes in flow experiences at work. *Hum. Relations.*, 65, 1101–1127. <https://doi.org/10.1177/0018726712447116>
29. Knierim, M. T., Nadj, M., Hariharan, A., & Weinhardt, C. (2018). Flow neurophysiology in knowledge work: Electroencephalographic observations from two cognitive tasks. In: *5th International Conference on Physiological Computing Systems (PhyCS)* (pp. 42–53).
30. Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32, 365–387.
31. Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, 32, 158–172.
32. Fest, S., Kvaløy, O., Nieken, P., & Schöttner, A. (2021). How (not) to motivate online workers: Two controlled field experiments on leadership in the gig economy. *The Leadership Quarterly*, 32, 16.
33. Lotze, M., Erhard, K., Neumann, N., Eickhoff, S. B., & Langner, R. (2014). Neural correlates of verbal creativity: Differences in resting-state functional connectivity associated with expertise in creative writing. *Frontiers in Human Neuroscience*, 8, 516. <https://doi.org/10.3389/fnhum.2014.00516>
34. Knierim, M. T., Nadj, M., Hariharan, A., & Weinhardt, C. (2018). Flow neurophysiology in knowledge work: Electroencephalographic observations from two cognitive tasks. In *Proceedings of the 5th International Conference on Physiological Computing Systems (PhyCS)* (pp. 42–53).
35. Rissler, R., Nadj, M., Li, M. X., Loewe, N., Knierim, M. T., & Maedche, A. (2020). To be or not to be in flow at work: Physiological classification of flow using machine learning. *IEEE Transactions on Affective Computing*, 1–12.
36. Piferi, R. L., Kline, K. A., Younger, J., & Lawler, K. A. (2000). An alternative approach for achieving cardiovascular baseline: Viewing an aquatic video. *International Journal of Psychophysiology*, 37, 207–217.

37. Moller, A. C., Meier, B. P., & Wall, R. D. (2010). Developing an experimental induction of flow: Effortless action in the lab. In B. Bruya (Ed.), *Effortless attention* (pp. 191–204). <https://doi.org/10.7551/mitpress/9780262013840.003.0010>.
38. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–183.
39. Tams, S., Hill, K., Ortiz de Guinea, A., Thatcher, J., & Grover, V. (2014). NeuroIS—Alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. *Journal of the Association for Information Systems*, *15*, 723–753.
40. Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*, 49–59.
41. Davis, F. D. (1985). A technology acceptance model for empirically testing new end-user information systems: Theory and results.
42. Kartali, A., Janković, M. M., Gligorijević, I., Mijović, P., Mijović, B., & Leva, M. C. (2019). Real-time mental workload estimation using EEG. In: *International Symposium on Human Mental Workload: Models and Applications* (pp. 20–34).