

Memmert, Lucas; Soroko, Daria; Bittner, Eva

**Article — Published Version**

## From Effort Reduction to Effort Management: An Expectancy Theory Perspective on Professionals' Work Practices with Generative AI

Business & Information Systems Engineering

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Memmert, Lucas; Soroko, Daria; Bittner, Eva (2025) : From Effort Reduction to Effort Management: An Expectancy Theory Perspective on Professionals' Work Practices with Generative AI, Business & Information Systems Engineering, ISSN 1867-0202, Springer Fachmedien Wiesbaden GmbH, Wiesbaden, Vol. 67, Iss. 5, pp. 615-635, <https://doi.org/10.1007/s12599-025-00960-4>

This Version is available at:

<https://hdl.handle.net/10419/330612>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# From Effort Reduction to Effort Management: An Expectancy Theory Perspective on Professionals' Work Practices with Generative AI

Lucas Memmert · Daria Soroko · Eva Bittner

Received: 14 October 2024 / Accepted: 16 June 2025 / Published online: 6 August 2025  
© The Author(s) 2025

**Abstract** Generative Artificial Intelligence (GenAI) is adopted by knowledge workers to boost productivity, yet its specific characteristics such as probabilistic outputs and human-level content generation may change how professionals think about their effort. Prior literature has warned about unintended side effects of AI, but experiments on effort reduction when working with AI – which could threaten performance – reported mixed results. GenAI's rapid adoption combined with its specific characteristics make it critical and timely to clarify how GenAI influences knowledge workers' effort in professional settings. The qualitative study draws on 21 interviews with knowledge workers who frequently use GenAI for work. A directed content analysis, guided by expectancy theory and social loafing frameworks, revealed that most interviewees do not simply reduce effort, but rather strategically reallocate or even increase effort. They continuously learn to steer GenAI, viewing themselves as process administrators. The traditional group-based mechanisms of reduced effort or diffused responsibility do not seem to be directly transferable to human–GenAI dyads in professional settings. By revealing that GenAI reshapes the factors that influence effort rather than simply eroding motivation, providing a multifaceted view of effort investment beyond mere reduction, and highlighting the interplay between human relationships and GenAI-facilitated work, this research advances the discourse on human-(Gen)AI dynamics and

the unintended consequences of (Gen)AI. Recognizing these shifts when setting policies and expectations enables organizations to benefit from GenAI's potential while mitigating potential risks to performance.

**Keywords** Generative AI · GenAI · Human-AI collaboration · Social loafing · Smart loafing · Knowledge work

## 1 Introduction

Text-based generative artificial intelligence systems (GenAI) such as ChatGPT are a subset of artificial intelligence (AI) systems capable of producing content in the form of fluent text. While GenAI is being rapidly adopted (Hu 2023) and can enable significant performance improvements for knowledge workers (Dell'Acqua et al. 2023), adopting AI may come with negative side-effects (Retkowsky et al. 2024). Studies suggest that working with AI might cause humans to reduce their efforts (Liu et al. 2023; Stieglitz et al. 2022), potentially “resulting in less-than-optimal team performance” (Liu et al. 2023, p. 2), if the AI cannot fully compensate for the reduction in human effort.

According to expectancy theory (Vroom 1964), individuals adjust the amount of effort they put into a task based on their belief that this effort will yield desirable outcomes. When *working alone*, individuals calibrate their effort to the perceived likelihood that additional effort will improve performance. Working with GenAI as compared to working alone fundamentally alters the task setting. GenAI can rapidly produce text and even creative content at a comparable level to many humans (Haase and Hanel 2023), conquering aspects of tasks previously thought to be

---

Accepted after two revisions by the editors of the Special Issue.

---

L. Memmert (✉) · D. Soroko · E. Bittner  
Faculty of Mathematics, Informatics and Natural Sciences,  
Department of Informatics, University of Hamburg, Vogt-Kölln-  
Straße 30, 22527 Hamburg, Germany  
e-mail: lucas.memmert@uni-hamburg.de

in the realms of humans (Dellermann et al. 2019). According to expectancy theory (Vroom 1964), this change in the task setting can tempt individuals to consciously or unconsciously question the marginal benefit of additional personal effort, and potentially lead to a withdrawal of effort, if workers believe their input will not further improve an already proficient GenAI output (i.e., low expectancy).

When tasks are performed *collectively in human groups* that produce a single group result, individuals may also reduce their effort, e.g., because responsibility diffuses, or because their personal contribution seems less critical (Karau and Williams 1993; Pinsonneault et al. 1999). The reduction of individual effort when working collectively is a phenomenon known as social loafing in human groups, which can decrease group performance (Karau and Williams 1993; Latané et al. 1979). Social loafing has further been formalized as part of the collective effort model (Karau and Williams 1993). Prior literature raised the question if humans show similar effort reduction behavior – akin to social loafing in human groups – when working with AI (e.g., Cymek et al. 2023; Hamza et al. 2024; Liu et al. 2023; Onnasch and Panayotidis 2020; Rix and Hess 2022; Siemon et al. 2015; Stieglitz et al. 2022). Observations that humans can perceive technical systems as social entities (Nass and Moon 2000), and that – when working with an AI – they might shift responsibility to the AI (Henkenjohann and Trenz 2024; Stieglitz et al. 2022), suggest potential commonalities of human–human and human–AI dyads. Experiments in prior literature investigating non-GenAI systems, however, show mixed results, e.g., some concluding that social-loafing “occurs” when working with AI (Stieglitz et al. 2022, p. 759) while others find no effect (Siemon et al. 2015).

Based on the above, we argue that GenAI’s distinct capabilities and interaction modalities may lead humans to reconsider their effort when working with GenAI, which could lead to “less-than-optimal” performance or introduce risks. In the light of the rapid adoption of GenAI in professional contexts and the mixed results of prior laboratory studies it is, therefore, important to understand how knowledge workers’ effort changes when using GenAI in professional work contexts. We, therefore, pose the following research question: *How do knowledge workers manage their effort when working collectively with GenAI in a professional context?*

To answer this research question, we conduct a qualitative study applying a directed content analysis approach (Hsieh and Shannon 2005). Drawing on previous research into changes in effort and loafing-like behavior in humans when working with AI, we have developed a conceptual model incorporating constructs such as expectancy, perceived responsibility and work performance. We conduct

21 interviews with knowledge workers who use GenAI in professional settings. The constructs in our literature-based conceptual model serve as theoretical framework for analyzing the interviewees’ reported effort investment rationales. We summarize the findings within these categories into subcategories to enhance clarity on how GenAI affects effort investment.

Our results illustrate several ways in which GenAI changes humans and their behavior, revealing that professionals do not simply reduce their effort, but strategically invest and reallocate effort to increase efficiency both on task- and job-level. More specifically, we explore how known antecedents for (reducing) effort are (re)shaped when working with GenAI, i.e., how GenAI may change human’s perception of how (additional or reduced) effort translates into outcomes. For example, although for some tasks knowledge workers reduce their effort when perceiving a GenAI-based version to be sufficient (signaling reduced *expectancy*), they continuously learn to effectively steer GenAI (signaling high *expectancy*). Instead of an across-the-board erosion of motivation due to AI (Liu et al. 2023), this suggests that motivational antecedents are reshaped. We find that most workers feel fully responsible for the results created with GenAI, contrary to findings of earlier studies where participants shifted responsibility to the AI (e.g., Henkenjohann and Trenz 2024; Stieglitz et al. 2022). Moreover, in line with earlier studies, we also find reduced effort for specific (sub-)tasks (Stieglitz et al. 2022), but our analysis suggests that workers tend to reinvest the saved effort into other (sub-)tasks. Lastly, we show that working with GenAI may affect relationships with others and vice versa, e.g., when creating work results with GenAI changes how others view and reward such results. Our analysis suggests that while established (collective) effort and social loafing perspectives offer a helpful lens for understanding how decisions concerning effort are made, describing the knowledge workers’ behavior as ‘loafing’ might be a mischaracterization.

With our work, we contribute to the discourses on human-GenAI dynamics (Makarius et al. 2020; Memmert and Bittner 2022), on the transferability of knowledge from human groups to human-(Gen)AI settings (Rix and Hess 2022; Stieglitz et al. 2022), and on unintended consequences of introducing (Gen)AI (Memmert et al. 2023; Retkowsky et al. 2024). Specifically, we contribute an overview and testable propositions on how GenAI may reshape knowledge workers’ motivation and effort in professional settings and discuss the potential influence of contextual factors. Building on prior conceptual work (Stieglitz et al. 2022), we suggest a multi-dimensional, multi-level *effort investment perspective* to reflect the inherent specifics of working with GenAI in professional settings, such as active prompting, probabilistic output

requiring verification, and re-allocating effort to various (sub-)tasks. Our work has implications for knowledge workers, their managers and organizations, highlighting the need to acknowledge this new way of working when interpreting professionals' efforts and work results.

## 2 Background

### 2.1 Foundations of Effort Investment Research

Effort may be defined as the “conscious exertion of power or hard work” (Merriam-Webster 2024), but is best conceptualized as a multidimensional construct comprising the three facets of effort *intensity*, *persistence*, and *direction* (van Iddekinge et al. 2023); accordingly, van Iddekinge et al. (2023) define effort as “how hard workers try to perform their jobs, which includes where they devote their effort (direction), the amount of their effort (intensity), and how long they persevere in their effort (persistence)”. Notably, while intensity and persistence describe how work is performed, direction reflects what work is performed – a distinction informing our comprehensive conceptualization of effort for this study.

Expectancy theory (Vroom 1964) proposes that individuals adjust their effort based on their subjective *expectation* or *belief* that additional input will yield valued outcomes. This theory posits that motivation – and, by extension, effort investment – is determined by three components: (1) expectancy, the belief that additional effort will enhance performance (Effort → Performance); (2) instrumentality, the belief that improved performance will lead to better outcomes (Performance → Outcomes); and (3) valence, the desirability or value of those outcomes. Consequently, effort decisions hinge on task context and the individual perception of whether additional effort exertion is worthwhile.

The Collective Effort Model (Karau and Williams 1993) extends expectancy theory to group settings. Working with others instead of alone alters the task setting. In groups, performance is determined not only by the focal individual's effort but also by the contributions of other group members, and shared rewards may not align with personal goals. These contingencies in collective group work weaken the perceived link between one's effort and outcomes, altering motivations to invest effort. How such contingencies map onto GenAI work is examined later.

Working collectively means “individuals work in the real or imagined presence of others with whom they combine their inputs to form a single group product” (Karau and Williams 1993, p. 681). If humans reduce their effort when working collectively rather than individually, this phenomenon is referred to as social loafing (Karau and

Williams 1993). Reasons for social loafing include the perception that one's effort is less crucial for group success when working with more proficient group members (dispensability of effort) or a reduction in individual responsibility (diffused responsibility) (Latané et al. 1979; Pinsonneault et al. 1999). This effort reduction may be conscious, following a deliberate appraisal of the task setting, or unconsciously, driven by “preexisting effort scripts” (Karau and Williams 1993, p. 685). Importantly, individuals adjust their effort not based on the actual effort-outcome-relationship but according to their belief or perception thereof (Karau and Williams 1993). Interestingly, these changes in effort are not only observable in actual effort exerted, but already in the intention to exert effort (Hüffmeier et al. 2013). Factors such as group size modulate loafing risk, which can emerge even in dyads (Cymek 2018; Cymek and Manzey 2022).

Prior research applied expectancy theory (Vroom 1964) as a framework for understanding motivation across diverse contexts, and social loafing has been documented in both physical and mental work settings (Karau and Williams 1993). Operationalizing (or measuring) effort, however, is not straightforward. Previous studies have employed both direct measures (e.g., the number of calls made by a sales agent) and indirect approaches, such as self-ratings or ratings by peers and superiors (van Iddekinge et al. 2023). van Iddekinge et al. (2023) pointed out that while some studies have equated effort and performance, the two constructs are correlated, but distinct.

### 2.2 Effort Investment in Sociotechnical Systems

Text-based GenAI systems – pre-trained on large text corpora and engaging users conversationally (e.g., ChatGPT) – are rapidly diffusing across organizations, enhancing performance in knowledge-intensive tasks (Dell'Acqua et al. 2023). As GenAI becomes increasingly capable – producing creative ideas that rival human performance (Haase and Hanel 2023; Dellermann et al. 2019) – humans may perceive their own effort as less essential (i.e., perceived *dispensability of effort*). Although working with GenAI differs from working with another human, prior literature indicates that certain dynamics observed when working *collectively* with AI toward an *integrated work product*, can resemble those found in human groups (Memmert and Bittner 2024; Memmert et al. 2024; Rix and Hess 2022; Stieglitz et al. 2022). Introducing (a probabilistic) GenAI weakens the effort-outcome relationship, because the work result becomes contingent upon the GenAI (to produce relevant output), analogous to the contingencies outlined in the collective effort model (Karau and Williams 1993). Humans may view technical systems as social actors (Nass and Moon 2000) or

teammates (Liu et al. 2023; Rix 2022; Rix and Hess 2022; Seeber et al. 2020), specifically ChatGPT (Hamza et al. 2024), and studies suggest that humans shift responsibility to AI systems (i.e., *diffused responsibility*; Henkenjohann and Trenz 2024; Stieglitz et al. 2022). Moreover, with human and GenAI being integrated into single work product, human and GenAI contributions may be difficult to discern (Retkowsky et al. 2024). In human group social loafing literature (Karau and Williams 1993; Latané et al. 1979), the above factors are recognized as key antecedents of reduced individual effort (i.e. social loafing); and Rix and Hess (2022, p. 13) suggest that “due the intertwined nature of [human–machine teaming], it is harder to disentangle machine and human conduct, fostering the emergence of social loafing.”

Researchers suggested examining whether social loafing-like behavior occurs in dyads of humans and technical systems, involving “AI-like” systems (Siemon et al. 2015), virtual agents (Stieglitz et al. 2022), robots (Cymek and Manzey 2022), or other AI tools (Liu et al. 2023), thereby extending the social loafing concept beyond human groups. Findings are mixed: some report no effort reduction (Siemon et al. 2015), while others documented self-reported effort reduction, suggesting that as systems are increasingly seen as “legitimate and equal team members” (Stieglitz et al. 2022, p. 757), humans may reduce their own effort, leading Stieglitz et al. (2022, p. 759) to conclude that social loafing “occurs” when working with virtual agents. As for perceived responsibility, some studies find no difference (Cymek and Manzey 2022), whereas as others found a shift to the AI (Henkenjohann and Trenz 2024; Stieglitz et al. 2022). The mixed results may reflect measurement challenges for effort (van Iddekinge et al. 2023) and heterogeneity of technical systems, but render whether social loafing-like behavior occurs when working with GenAI uncertain. Moreover, whether social loafing is the right term is debated, particularly when AI compensates for reduced effort, leading Stieglitz et al. (2022, p. 758) to reframe it as “smart loafing,” where humans strategically reduce effort to conserve cognitive resources and enhance work efficiency.

Even if not fully subscribing to the view that GenAI induces social loafing per se, expectancy theory (Vroom 1964) predicts that any shift in task structure – such as the integration of GenAI – will recalibrate individuals’ subjective evaluations of whether additional effort is worthwhile. Because GenAI can take over facets of knowledge work, users may question whether their contributions meaningfully improve outcomes and thus may adjust – or even reduce – their effort when working with GenAI. Importantly, reduced human effort does not necessarily translate into lower overall performance: GenAI’s efficiencies may offset or even surpass baseline outputs.

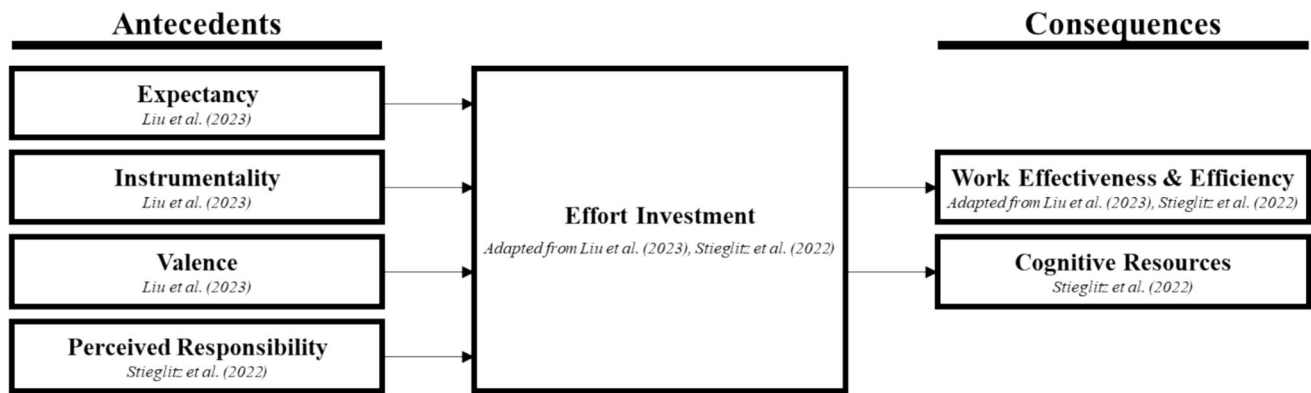
However, reduced human effort can leave AI potential untapped or yield “less-than-optimal team performance” (Liu et al. 2023, p. 2). Accordingly, we investigate how GenAI shapes knowledge workers’ effort investment and revisit whether social loafing is an appropriate lens for GenAI work.

### 3 Derivation of the Conceptual Model

Decades of research on both expectancy theory (Vroom 1964) and social loafing (Latané et al. 1979) across domains and across both mental and physical work demonstrates that individuals adjust their effort based on subjective perceptions (van Iddekinge et al. 2023; Karau and Williams 1993). As GenAI becomes increasingly capable and may be perceived as partners when performing knowledge work collectively (Hamza et al. 2024; Rix and Hess 2022), they may change workers’ perceptions of the impact of their own effort. To analyze how GenAI might shape knowledge workers’ effort investment decisions – directly and through its antecedents – we propose a conceptual model as a theoretical framework (see Fig. 1), which we have derived from literature as described below.

Literature on expectancy theory and social loafing is vast (van Iddekinge et al. 2023). To ensure relevance with respect to the emerging discourse on social loafing-like behavior and effort management when humans work with technical systems, we derive our model from the recent works of Stieglitz et al. (2022) and Liu et al. (2023), as these studies, too, discuss changes in effort investment with conversational or AI systems in one-on-one settings (i.e., one human working with one system) when performing knowledge work. From these papers, we extract the discussed antecedents as well as the consequences of effort reduction when working with AI. We include factors that are relevant to effort investment and consolidate and adjust them for conceptual clarity. We exclude fixed personal traits (e.g., personality factors) that are less likely to immediately shift when working with GenAI, as we sought to understand the impact of GenAI on knowledge workers’ effort investment – though we acknowledge that personal differences likely affect those decisions. Each construct is briefly discussed below and defined as a foundation for our analysis of the interviews with knowledge workers. While for the selection of the antecedents and consequences (e.g., defining the scope of the analysis), we used the works of Stieglitz et al. (2022) and Liu et al. (2023) as a foundation to ensure relevance to the emergent discourse on collective human-AI knowledge work (see rationale described above); for the definition and explanations of those constructs, we used literature on expectancy theory and social





**Fig. 1** Conceptual model derived from Liu et al. (2023) and Stieglitz et al. (2022)

loading more broadly, to ensure consistency with prior literature.

For antecedents – following Liu et al. (2023) – we include *expectancy*, *instrumentality*, and *valence*, strongly rooted in expectancy theory and the collective effort model (Karau and Williams 1993). Expectancy refers to the belief that additional effort will enhance performance; instrumentality to the belief that improved performance will lead to better outcomes; and valence to the desirability or value the human places on these outcomes (Karau and Williams 1993). Additionally, we include *perceived responsibility* for the work outcomes as an antecedent following Stieglitz et al. (2022), who discussed the attribution of responsibility between the human and the technical system. We did not explicitly include expertise (as mentioned by Stieglitz et al. 2022), as this is subsumed under expectancy, i.e., the expectation of the degree of personal capability compared to the AI capability (also see discussion in Liu et al. 2023).

The central construct in our conceptual model is effort investment. While Stieglitz et al. (2022, p. 761) operationalized human effort via “social loafing tendencies in virtual collaboration with the virtual assistants”, i.e., with an effort-reduction-perspective, we adopted a more neutral perspective on effort, more akin to the conceptualization of Liu et al. (2023, p. 5), who referred to it as “human effort in human-AI team.”

For consequences of changed effort investment, we included work effectiveness and efficiency as well as cognitive resources, as these are the central concepts in the “smart loafing” construct developed by Stieglitz et al. (2022, p. 758), which they define as “the reduction of effort in human-[virtual assistants] collaboration to maintain *cognitive resources* and enhance *efficiency in work*” (emphasis added). *Work effectiveness and efficiency* refers to how the work process and outcomes are affected by the changed effort. Following the smart loafing perspective (Stieglitz et al. 2022), for *cognitive resources* we refer to the personal mental capacity being conserved or depleted

while working on the task, e.g., due to increased or reduced cognitive load (Sweller 1988). Though cognitive resources could be included into efficiency, we decided to differentiate the two constructs due to the difference in the referend, i.e., task versus individual.

We do not challenge the underlying theories, which have been tested across many domains and contexts. With our model, we integrate a sub-set of antecedents and consequences contained in the model of Liu et al. (2023) and the smart loafing definition of Stieglitz et al. (2022), to guide the investigation of how GenAI shapes motivation and effort investment decisions of knowledge workers in organizations (Table 1).

#### 4 Method

To investigate how professionals determine their effort investment when using GenAI in their daily work, we used a qualitative approach, interviewing 21 knowledge workers. The interviews took place between March 2024 and March 2025. The interviewees were recruited via a professional society newsletter as well as the authors’ personal network. We selected only employed professionals to understand their effort investment decisions in a work context embedded in organizational structures, where the importance of the task outcome is representative for work settings and matters to other stakeholders. When selecting participants, we ensured that all interviewees met the established criteria for knowledge work, including white-collar roles (Mandt 1978, as cited in Timonen and Paloheimo 2008), low standardization (Pyöriä 2005), engagement with abstract knowledge and symbols (Pyöriä 2005), reliance on specialized skills and theoretical knowledge (Timonen and Paloheimo 2008), non-routine tasks (Järvenpää and Eloranta 2001, as cited in Timonen and Paloheimo 2008), problem-solving abilities (Tsoukas and Vladimirou 2001, as cited in Timonen and Paloheimo

**Table 1** Literature-based category system used as deductive frame with categories and explanations

Category	Explanation	References
Expectancy	Refers to thoughts about the believe of the ability to improve task performance, as lower expectancy can lead to effort reduction, whereas the feeling of being capable to make a unique contribution might make effort reduction less likely to occur	Karau and Williams (1993); Liu et al. (2023)
Instrumentality	Refers to thoughts on the necessity to contribute to achieve better outcomes through improved performance, as lower perceived instrumentality (i.e., a higher perceived dispensability of one's effort) can lead to a reduction in effort	Karau and Williams (1993); Pinsonneault et al. (1999); Liu et al. (2023); Memmert and Tavanapour (2023)
Valence	Refers to thoughts on how valuable or desirable the knowledge worker perceives the outcome of the tasks to be, with high valence being associated with a reduced likelihood of reducing effort	Liu et al. (2023); Williams and Karau (1991); Karau and Williams (1993)
Perceived responsibility	Refers to thoughts on the responsibility for work results, as (perceived) responsibility diffusion might lead to a reduction in effort (responsibility attribution towards technical systems occurred in prior literature)	Pinsonneault et al. (1999); Stieglitz et al. (2022); Henkenjohann and Trenz (2024); Memmert (2024)
Effort investment	Refers to thoughts about how to calibrate the own effort invested in the task, excluding general strategies on how to work effectively with GenAI, like prompt engineering	Stieglitz et al. (2022); Memmert (2024)
Work Effectiveness and Efficiency	Refers to thoughts and the reasoning on leveraging GenAI for the tasks, as enhancing work efficiency was highlighted as a key reasoning for leveraging technical systems	Stieglitz et al. (2022); Liu et al. (2023)
Cognitive Resources	Refers to thoughts about the management of and effect on mental resources when working with GenAI	Sweller (1988); Stieglitz et al. (2022); Hart and Staveland (1988)

2008), and the need for extensive formal education and continuous on-the-job learning (Pyöriä 2005). Our sample also included knowledge workers who had experience using GenAI, such as ChatGPT, in their daily work. We used a broader sample to get a more general understanding of how knowledge workers think about the effort investment and not be biased due to a specific type of profession. We have listed the industries, job titles and overall professional experience in Table 2.

The interviews were semi-structured (Meuser and Nagel 2009), following an interview guideline developed along our research goal and conceptual model (Fig. 1), but leaving room for personal experiences of the professionals. The guideline underwent a slight revision after the first 16 interviews to incorporate more detailed clarifying questions (see Appendix A, available online via <http://link.springer.com>). After providing a general introduction to the study and obtaining consent, we began the first part of the interview by asking about interviewees GenAI usage in professional work, and – while not at the core of our study – about the tasks they use GenAI for. This allowed us to build rapport with the interviewees, preparing them for sharing more detailed reflections in the interview, but also allowed us to contextualize the results. In the second part, we asked participants to reflect on their effort investment when using GenAI. Depending on the answers, we asked follow-up questions, e.g., regarding the interviewees' first steps when working jointly with GenAI on a task, routines

they had developed while using GenAI, how they determined when to stop working together with GenAI on the task, and how they integrated the outputs into their final work results. In the third part, we asked participants to reflect on how the work results produced with GenAI felt to them personally, and how they felt working with GenAI affected their mental workload and stress. Interviews were closed with an open-ended question on whether participants wanted to share any additional information. We made sure to ask open-ended questions as well as indirect questions to prevent biasing the respondents (Döring and Bortz 2016). To make the interviews concrete, we asked participants to share specific examples (if they were allowed to disclose the information) or even share their screen to walk us through their conversations with GenAI. We used this approach to invite participants to reflect on their behavior, because – as was pointed out in the background section – humans may sometimes adjust their effort unconsciously within certain situations and might become only aware when actively reflecting on it.

All interviews were conducted online, were recorded with the participants' permission, and were later transcribed verbatim. We followed a qualitative content analysis approach (Hsieh and Shannon 2005; Mayring 2014), more specifically, a directed content analysis (Hsieh and Shannon 2005). Hsieh and Shannon (2005, p. 1281, emphasize added) explain that the “goal of a directed approach to content analysis is to [...] *extend conceptually*

**Table 2** Interviewee details

ID	Job	Professional Area	Professional experience range [years]	Interview length [mins.]
I1	Manager	Energy Infrastructure	5..8	38
I2			5..8	33
I3			5..8	41
I4		Consulting	9..12	46
I5			1..4	39
I6			9..12	31
I7	Leader and Manager	Industry Park	9..12	35
I8			5..8	32
I9			1..4	45
I10	Consultant	Consulting	5..8	39
I11			> 12	42
I12			1..4	42
I13	Associate or senior researcher	Research	5..8	44
I14			5..8	65
I15			5..8	58
I16			> 12	48
I17			1..4	33
I18			> 12	52
I19			5..8	63
I20			5..8	50
I21		Automotive	5..8	33

a *theoretical framework* [...]”, which is helpful if “*existing theory or prior research exists* about a phenomenon that is incomplete or *would benefit from further description*”. This approach fits well with our research goal, because established theories *exist* (i.e., expectancy theory, collective effort model) explaining effort changes in human groups, which underwent initial extension to work with AI-based systems and were integrated into a *theoretical framework* (see Fig. 1). We now seek to explore and *describe* how human effort changes when working with GenAI. Following Hsieh and Shannon (2005, p. 1281) we begin by using “existing theory or prior research” to identify “key concepts [...] as initial coding categories” to guide our analysis and add “operational definitions for each category”; see Fig. 1 with detailed explanation in Sects. 2 and 3, and Table 1 with the categories and definitions. We then proceed to code the interviews (appendix B contains the full coding schema). Individual statements with complete arguments were used as primary coding units. As suggested by Hsieh and Shannon (2005), we grouped our findings into sub-categories to enhance clarity. Two interviewers first coded a small sample of all interviews (two interviews each) to shape the initial coding scheme. This was followed by a round of discussions to resolve any disagreements and inconsistencies between the coders. Then the same sample

of interviews was re-coded individually once more to align data analysis to the revised coding scheme. Finally, the rest of the interviews were analyzed according to the established coding guidelines. To make the results more tangible, we selected direct quotes from the interviews, which were translated for interviews not conducted in English, to be included in the final manuscript.

After coding the initial set of 16 interviews, we reached data saturation, defined as the point at which no new (sub-)categories or insights were emerging from the collected data. Saturation was assessed by both coders through ongoing discussions throughout the data analysis by reviewing transcripts and developed (sub-)categories to ensure that they adequately captured the data. The subsequent five interviews did not yield any new information and confirmed earlier findings. This is consistent with prior research suggesting that data saturation in qualitative studies can be reached as soon as twelve interviews (Guest et al. 2006).

## 5 Results

To contextualize the interviewees’ reflection on working with GenAI, we first asked them about which GenAI they



used, and how frequently and for what kind of tasks they used GenAI. We find the interviewees mostly use OpenAI's ChatGPT and some use Microsoft Copilot. Many interviewees use GenAI multiple times daily and all use GenAI for their professional knowledge work. This was expected given our research goal and sampling criteria. Our interviewees cover a variety of tasks, from low- to high-stakes and from operational to strategic tasks. Examples include summarizing notes (e.g., I10), formulating, grammar checking, or editing emails and texts (e.g., I12–I15), translating documents (e.g., I9–I10, I14), sanity-checking concepts or presentations (e.g., I8, I21), developing user stories (e.g., I10–I11), developing training material and exercises for students or fellow employees (e.g., I6, I7, I10, I14, I16) or developing concepts (e.g., I9, I11, I18). Most of the tasks met our knowledge work criteria mentioned in chapter 4. The reported “conversation lengths” with the GenAI (i.e., number of messages sent by the user) differed from one-off interactions up to week-long conversations.

We have structured the results section according to the conceptual model presented earlier, with effort investment, and potential antecedents and consequences. We have summarized the interviewees' considerations reported in Fig. 2, which may increase, retain, decrease certain model elements; we have indicated the *tendencies* according to the interview data, or flagged them as ambiguous. However, these need to be treated as preliminary and with caution, requiring (quantitative) validation.

### 5.1 Expectancy

Expectancy refers to the human's belief that (additional) effort will improve performance (Vroom 1964). Our results show that – when working with GenAI – this can involve not only the human's own direct contributions but also their perceived ability to steer the GenAI towards creating better outputs, and even enablement to work on tasks that the human could not have performed alone.

*Own contribution* (→) Generally, participants' responses reflect a high sense of expectancy with most participants reporting that they believe they can make own, direct contributions that improve performance. Specifically, many participants noted that they provide overall *strategic or structural knowledge* for the task (e.g., I11, I18–I20). Additionally, they report that they provide the *situation-specific, expert knowledge* (e.g., I17, I18–I20) required for most tasks, that cannot be generated by GenAI, with one participant explaining “the results [produced by the GenAI] are semi usable if you don't put your own know-how into it [...]. What we do requires a certain creative

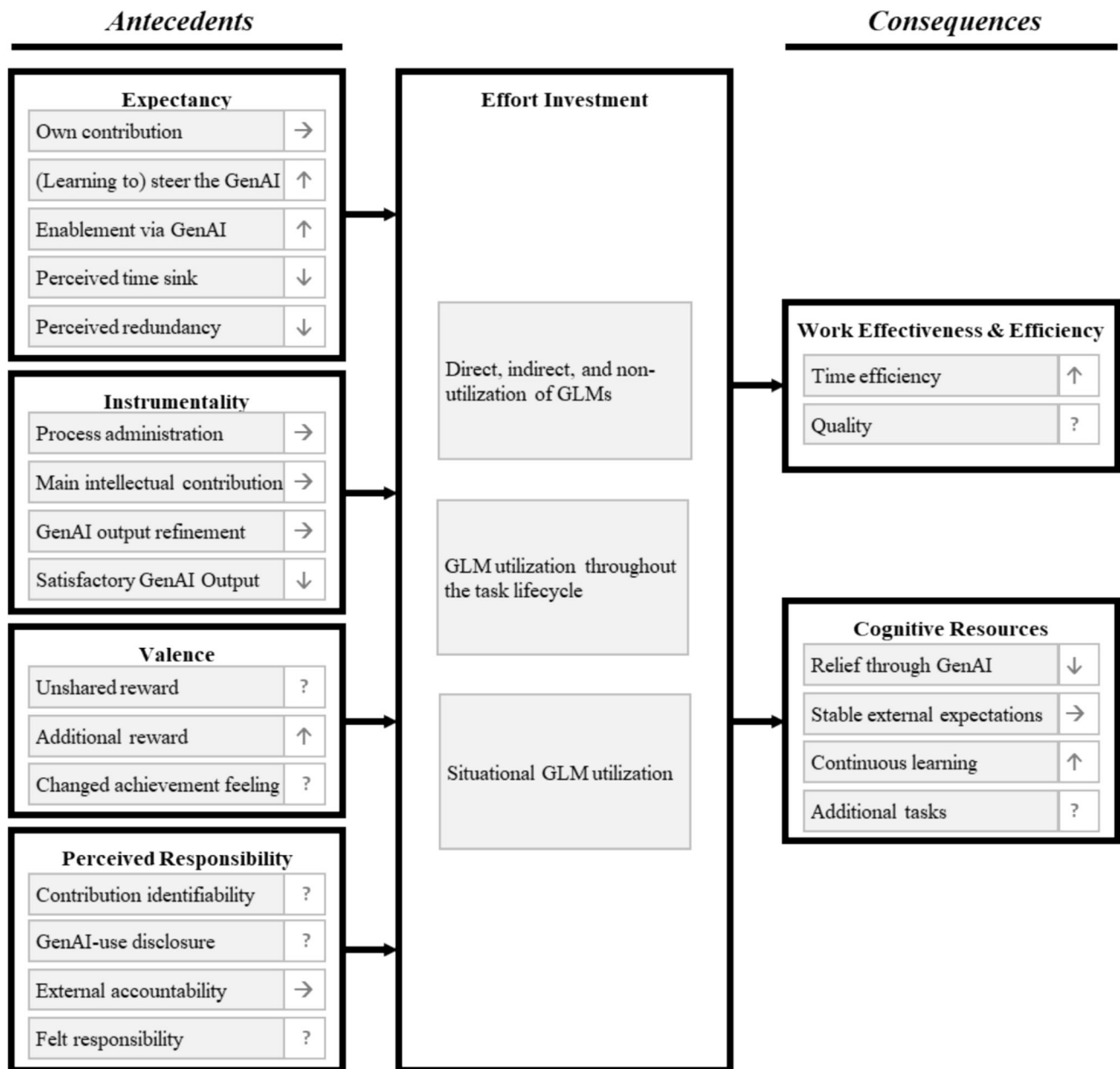
approach, which the systems currently do not yet provide” (I4).

*(Learning to) steer the GenAI* (↑) Beyond making direct contributions, participants also reported investing up-front and continuously into learning to effectively steer GenAI. This includes, e.g., reading blog articles and prompting guides as well as reviewing co-workers' conversations with the GenAI (e.g., I11) and attending workshops (e.g., I16). One interviewee (I5) explained “especially at the beginning, as I said, I looked a lot at blog articles and things like that on how you can, may, should optimize your workflows with [GenAI]; what kind of prompt is suitable or whatever. But now I think I know that I at least have a set of things that work quite well for me.”

This covers operational, conceptual, and procedural aspects. On an operational level, interviewees try to learn how to effectively prompt, e.g., deliberately providing context information, certain perspectives or keywords, specifying guardrails or frameworks to follow (e.g., I9, I19). They calibrate the prompt complexity for the task complexity (e.g., I3, I9). Beyond prompting design techniques, some learn and re-use prompts for recurring tasks (e.g., I10, I13–I16) or have even developed prompts library (e.g., I3, I15). Conceptually, interviewees test and try to understand the boundaries of GenAI (e.g., I9, I16, I20), and adjust their usage behavior accordingly, e.g., by editing prior messages in conversations to remove unneeded information (e.g., I9), or by fact-checking GenAI outputs. One participant (I20) stressed the importance of having an idea for how GenAI functions to be able to delegate tasks appropriately. On a procedural level, the focus lies on understanding how to integrate GenAI into their own workflow (e.g., I5, I15, I19).

Learning how to effectively utilize GenAI is not perceived as a one-time effort, but as a continuous learning activity (e.g., I7–I8, I11, I15–I16), requiring effort to improve prompting efficiency, advancing the prompt library, and monitoring the evolving GenAI tool landscape with different GenAI systems (e.g., I9, I10, I14). This sustained effort investment illustrates participants' belief that in-situ steering or learning to better steer the GenAI leads to greater control over the model's output and, in turn, meaningfully improves performance. This belief of exerting effort in the interaction or even investing effort up-front to increase performance, signals a high level of expectancy (Vroom 1964).

*Enablement via GenAI* (↑) Participants report to be enabled to perform new tasks they otherwise could not have performed. Thus, while participants might have felt low expectancy with respect to certain tasks pre-GenAI,



**Fig. 2** Changes to knowledge workers' effort investment when working collectively with GenAI.

Note: We assessed the reported influence of working with GenAI on antecedents and consequences of effort investment; arrows indicate the reported direction – increase (↑), retained (→), decreased (↓), or ambiguous (?); direction should be interpreted with caution and require further (quantitative) validation

due to GenAI, their skills are elevated making their effort productive towards achieving new outcomes, e.g., working on a task with aspects outside their expertise (e.g., I6, I17–I18). An interviewee recently started learning to code, and with the help of ChatGPT they feel capable of tackling most of the arising coding problems (I9). Others explained, the GenAI allows them to focus on the structuring and creative part of the task (e.g., I5), or that they felt “powerful” when having access to ChatGPT: “At the fingertips, I have so much power now that I can use!” (I12).

*Perceived time sink* (↓) Signs for reduced expectancy were rare. While most participants believe they can (learn to) control GenAI, some participants pointed out that using a GenAI is sometimes perceived as a time sink, when they cannot get the GenAI to produce the output they need for certain tasks (e.g., I16), with one participant explaining “I spend so much time now essentially arguing with GPT about a solution that I’m not happy with. And this was really not fruitful in any way and kind of a waste of my time” (I14). This shows a parallel to the contingency in

human groups discussed by Karau and Williams (1993): when working with others, the relationship between own effort (input) and the work result (outcome) is less direct (as compared to when working individually), because the outcome is dependent on additional factors such as other group members' contributions. Similarly, when working with a GenAI, the outcome becomes contingent upon the GenAI to produce relevant output. When professionals feel that interacting with GenAI burns time without improving its output, their expectancy – the belief that extra effort will pay off – is reduced. Overall, GenAI is not viewed a waste of time, but as a means to enhance productivity. Interestingly, several participants perceived working with GenAI as positive even when not receiving the expected results, using it as a learning opportunity.

*Perceived redundancy* (↓) However, some participants pointed out that for some parts of their task like phrasing or expressing thoughts in another language, they perceive the GenAI as more proficient, rendering their contribution – for this aspect of the task – redundant or irrelevant (e.g., I17), signaling reduced expectancy.

## 5.2 Instrumentality

While expectancy focusses on whether additional effort will improve performance, instrumentality refers to the belief that improved performance leads to better outcomes (Karau and Williams 1993; Vroom 1964). When group members feel that their effort is less instrumental (i.e., feel their effort to be dispensable) for group success, they might reduce their own effort. In the prior subsection, it was reported that participants generally felt that despite GenAI becoming more capable, participants' effort would lead to enhanced performance (expectancy). Beyond expectancy, most participants felt that their contribution not only enhanced performance but was *necessary* to achieve the work results that stakeholders expect. Hence, their effort is *instrumental* in achieving better outcomes, which indicates sustained instrumentality despite the GenAI. Participants explained this with respect to the way GenAI is integrated into their job more fundamentally (see *process administration*), the nature of participants' contributions (see *main intellectual contribution*), and the GenAI's output quality. Only in rare cases was the GenAI's output quality sufficient on its own, indicating a potential reduction in instrumentality.

*Process administration* (→) We asked participants to reflect on how they integrate GenAI into their work. Participants reported that they understand the task and expectations, enter the required information into the GenAI, iteratively develop, evaluate and refine GenAI outputs, and shape them according to stakeholder expectations. Because participants still orchestrate the workflow

and integrate the knowledge into the organization – leaving the core of their work untouched (I5) – their effort remains indispensable for achieving the desired outcome, signaling unchanged instrumentality.

*Main intellectual contribution* (→) On a more detailed level, participants point out that usually it is them who contribute the main intellectual, structural, or strategic knowledge, while GenAI serves more as an editor (e.g., I13, I17). Hence, while GenAI can support with many task aspects, without the participants there would not be a sufficient foundation for the GenAI to produce work results in the first place. Interviewees felt able to make unique contributions, e.g., in the form of their expert knowledge and experience (I10, I12–I14, I19–I21), whereas unique contributions by the GenAI were rare (e.g., I6, I11, I19). Supplying the strategic knowledge that the GenAI lacks keeps human effort instrumental to the final outcome.

*GenAI output refinement* (→) On a more operational level, participants state that they cannot use the output as-is, explaining that it is usually too general or contains false information, resulting in them having to fact-check, refine, and enrich the output with their own expert knowledge (e.g., I5, I7–I8, I19–I20). Interviewees assure quality with varying degrees of effort investment by them and others. In more general terms, interviewees stated to have a certain quality-related “self-image” of their work (e.g., I6, I9, I13–I14), both personally and for their company towards their clients, or as one interviewee (I5) put it “It's also somewhat of a principle of our work to only send out high-quality results”. When engaging with the GenAI to work on a task, interviewees reported to have developed expectations regarding the results upfront (e.g. I2–I3, I6, I18–I19). Given participants generally seem to be aware of (some of) the limitations of GenAI – e.g., naming hallucinations (e.g., I3) – they report to critically review and edit the GenAI's output based on their own (expert) knowledge and experience (e.g., I4, I7–I8, I12–I15, I19–I21), while also considering stakeholder expectations (e.g., I2), and to verify the outputs with original sources (e.g., I5, I18–I20); or, as one interviewee (I3) put it: “In my opinion, we will never be spared the step of saying that a human who knows what they're doing needs to look at [the GenAI output] again to say whether it fits or not.” From these explanations it becomes clear that fact-checking and enrichment is perceived to be necessary, meaning the task cannot be completed without human intervention, sustaining instrumentality.

While stressing having checked the GenAI's output, interviewees revealed instances in which they partially outsourced quality assurance. One interviewee (I6) shared that they had to create a document in a field outside their expertise. Given this was a one-time task and familiarizing oneself with the relevant literature would have meant a

large effort investment, the interviewee instead leveraged a GenAI to produce the document and sent it to a different department, which holds the final accountability. A different approach that was mentioned was to clearly disclose to have used GenAI to manage expectations that the output might be taken with a grain of salt, or might not be helpful (I2, I10).

*Satisfactory GenAI output (↓)* While participants generally indicated a high sense of instrumentality and usually refine GenAI-produced output, several participants pointed out that for some tasks as part of their jobs in the organization the GenAI output or “80%-solutions” are sufficient (e.g., I2). This means that the humans’ effort would not be required anymore for those tasks because – even though humans could improve performance (expectancy) – such increased performance would not lead to better outcomes (because GenAI outputs already meet organisational standards), indicating a reduction in instrumentality for these aspects of their job.

### 5.3 Valence

Valence refers to how valuable the human finds the outcome when completing a task (Karau and Williams 1993; Vroom 1964). Compared to individual work, the connection between individual performance and individual outcomes are less direct in group settings due to additional contingencies (Karau and Williams 1993). Rewards may be provided to the group but may not necessarily translate into valued outcomes for the individual, e.g., because the group gets praised, but the individual’s contribution is not recognized. We were curious if working with GenAI affected valence, i.e., if working with GenAI changed the value perception of the work results. However, results show that generally, participants feel similarly rewarded even when working with GenAI and might even receive more rewards. However, working with the GenAI can alter the feeling of achievement.

*Unshared reward (?)* With GenAI partially taking over, it might be conceivable that others, when evaluating the work results, reduce their praise or credit, because not everything was done by the human alone, analogous to human group settings, where praise may be divided among group members. However, participants reported that they generally do not face this problem (e.g., I18). Work results are usually perceived as theirs, hence, they receive the reward (i.e., unshared reward), indicating sustained valence. However, there were instances in which superiors joked about the contribution of the human (e.g., I10), given that a large part of the results was produced with GenAI. Thus, with even more widespread adoption, peers and managers might indeed start judging work results differently when they can be produced (to a large extent) with

GenAI, potentially deferring praise partly to “the GenAI,” and reducing praise for the involved human.

*Additional reward (↑)* Additionally, participants pointed out that because they can perform tasks better; focus on more value-adding aspects of the tasks (e.g., I5), can perform more work in general being more productive (e.g., I5), and are even enabled by GenAI to do tasks they could have not done without it (e.g., I9), i.e., they can do more and potentially received additional rewards they might have not received otherwise.

*Changed achievement feeling (?)* Beyond the extrinsic, there is an intrinsic dimension to valence (Karau and Williams 1993). GenAI may affect how participants feel about (some of their) work, i.e., when developed jointly with GenAI, they perceive it as “less” or have a reduced feeling of accomplishment (e.g., I2, I9), which could reduce intrinsic motivation. One participant (I9) explained it as follows “It feels a bit like someone else is solving the problem for you.” However, most participants explained that they weave the GenAI’s output into the work results (e.g., I18), reporting to feel it to be “their” work, e.g., stating “so after rewriting it a few times, I think it has become my text rather than a generated text” (I16). This point is closely connected to a feeling of responsibility (next section). One participant (I19) also reported creating a very particular routine when using ChatGPT for most of her tasks, which enabled her to “actually enhance the meaningfulness of the task” and made her feel much happier about the outcome.

### 5.4 Perceived Responsibility

An important factor influencing the level of effort is how responsible a person feels for the task and outcome (Karau and Williams 1993; Pinsonneault et al. 1999). Literature on social loafing shows that making individual contributions in group work identifiable, can increase responsibility and reduce social loafing behavior (Karau and Williams 1993). Whether or not recipients of the work results can identify the parts of the work results produced by the human versus by GenAI, depends on whether (1) they can tell the GenAI and human contribution apart, or, if they cannot, on whether (2) the person sharing the work results tells the recipients if and how they have used GenAI when producing the results. We inquired participants on both questions. Additionally, as prior literature showed that humans might shift responsibility to AI (Stieglitz et al. 2022), we also asked for the felt responsibility more directly.

*Contribution identifiability (?)* Responses on whether interviewees felt that the recipients of their work results can distinguish, which parts originated from them and which parts from GenAI were mixed, some said the

contributions are indistinguishable (e.g., I3, I6–I7, I16, I20), e.g., because they worked iteratively and integrated the GenAI's and their own thoughts closely in the process. Other participants were unsure (e.g., I5), or believed contributions to be distinguishable (e.g., I13–I15). Hence, the identifiability of contributions may depend on contextual aspects like usage behavior rendering the effect on perceived responsibility uncertain.

*GenAI-use disclosure (?)* When interviewees have the feeling their contributions and those by the GenAI are not distinguishable for the recipient, they have the choice of disclosing having worked with GenAI or not. Generally, interviewees seem not to disclose having worked with GenAI (e.g., I3, I5–I6, I12–I14, I20–I21), drawing a comparison with other tools like Google, which they also do not explicitly disclose to have used (e.g., I3, I8); explaining that “the part they generate is usually only a portion of what [they] send out” (e.g., I5), with the majority being written by them; that there is no organizational rule requiring such disclosure (e.g., I3); or that they also do not “explicitly claim it as [their] result” (I6). In contrast, reasons for disclosing having worked with the GenAI include promoting a new way of working and encouraging using GenAI for work (e.g., I10); working closely together with people (e.g., I6); or certain rules requiring disclosure (e.g., I12–I15, I21). Additionally, a reason for disclosure might be not wanting to take false credit and managing expectations towards co-workers regarding one's ability, or as one interviewee put it: “Why should I claim something as mine, if it is not and something I don't do it well?” (e.g., I10). Lastly, as mentioned above, disclosure might also be used to help co-workers to better assess the quality of the results (e.g., I10). Choosing whether – and how – to reveal GenAI use gives workers control over identifiability, so responsibility effects may vary.

*External accountability (→)* After discussing known causes associated with reduced responsibility and social loafing behavior (Pinsonneault et al. 1999), we directly about participant's felt responsibility and ownership regarding the results: most if not all participants felt highly responsible. Reasons mentioned included being held accountable for the results and being judged by the outcomes of their work with expectations remaining the same (e.g., I6, I19) irrespective of how results are produced, or, as one interviewee put it: “nobody cares about the way the results are produced – it is perceived as coming from me” (I2).

These results surface an interesting difference between established knowledge on human groups and the human-GenAI dyads in our study when it comes to identifiability of contributions. For recipients of the work, it may not be easy to understand, which parts of the work results were actually produced by the human (and which by GenAI),

and the creator of the work result has a choice of disclosing whether and how they used GenAI. Based on prior literature, non-identifiability and non-disclosure together could mean, that the actual contribution of the human in concealed, which might allow the human to reduce their effort, analogous to social loafing behavior. However, unlike in human groups, there are no other humans when working with GenAI. Because stakeholders treat the human as the sole producer, participants reported being held as responsible for the outcome as if they had not used the GenAI.

*Felt responsibility (?)* Regarding whether their work results produced alone felt different as opposed to when developed using GenAI, responses among interviewees were mixed. Some interviewees stressed they feel results produced with GenAI are their work (e.g., I3, I6, I12–I16, I20–I21), offering varying reasoning, e.g., that the main thoughts for the result still come from them (e.g., I11–I15, I18, I21), that they had written the prompts (e.g., I3, I11), or that they had invested time (e.g., I14). One of the most prominent points made was a sense of intellectual or creative ownership, with several participants stating that the main ideas and intellectual effort came from them (e.g., I17–I18), while GenAI was more of an editor or assistant tool (e.g., I12–I15, I21). Thus, being the main intellectual contributor and process orchestrator does not only lead to instrumentality (as reported earlier), but also to feeling responsible.

An additional perspective was provided by one interviewee (I6), who explained that being a manager they delegate many tasks anyways, so using GenAI would not change how they perceived the work results, stating “it is about the task” (I6). Other interviewees, however, reported the results produced with GenAI felt different. Potential explanations offered included having to spend less time developing the results, and, thus, having felt less personal attachment (e.g., I1); that the results sounded more eloquent than the interviewee could have formulated it (e.g., I10), and that working with GenAI is a bit like taking a short-cut (e.g., I2) or “it feels a bit like someone else is solving the problem for oneself” (I9), which takes away from the feeling of achievement (e.g., I2, I9). This does not necessarily mean they deny responsibility, but they might feel less personal attachment. Hence, the felt responsibility may vary, e.g., depending on personal factors (such as how one views one own work) and on how GenAI is integrated into the work.

## 5.5 Effort Investment

We asked interviewees to reflect on their considerations towards managing their effort investment when working with GenAI, finding participants have developed certain



patterns on when, how, and why they use GenAI, with implications for effort intensity, persistence, and direction.

*Direct, indirect, and non-utilization of GenAI.* Several interviewees stated to include large portions of GenAI's output in the final work results, with estimates ranging from 50 to 85% of text copied (e.g., I3, I5–I6, I14), particularly for texts in which GenAI is used as a writing assistant. Against this backdrop, a participant (I17) reported to sometimes use a simpler writing style just to get something on paper, because they know GenAI will rephrase appropriately, explaining “before you don't write anything because you're thinking too much, I've gotten into the habit of just starting to write and then I can always read about it later or just tell the AI now, can you please go over my low effort here and work it up a bit linguistically”. Interviewees not only use the system directly for their work result, but also indirectly, e.g., to produce speaker notes for slides (e.g., I2), to get a better understanding of a topic (e.g., I5, I9, I15, I19–I20), or for internal knowledge organization more generally (e.g., I9). However, though helpful for many tasks, GenAI is not capable to support many other tasks in the interviewees' jobs (e.g., I7–I8, I16) – interviewees deliberately choose for which tasks to engage GenAI, and when to perform the task alone.

*GenAI utilization throughout the task lifecycle.* Social loafing-like behavior might result in humans reducing their effort to let GenAI do their work. We, thus, asked the interviewees about how they used GenAI throughout the task lifecycle. Interviewees seem to employ GenAI throughout working on their tasks, from the very beginning, throughout performing the task, and in the end, with varying reasoning. In the beginning, interviewees use it as a replacement for doing their own research, to lower the entry barrier and quickly get an overview of a topic outside their expertise (e.g., I5, I7, I9–I10, I15, I18–I20), to quickly generate draft versions with just a few inputs (I5) instead of starting with a “blank paper” (e.g., I1, I14, I8), or to generally get inspiration on “where to go from here” (e.g., I6, I12–I13). When used throughout working on a task, interviewees use GenAI to get a different perspective (e.g., I10, I21) or as a “sparring partner” (e.g., I3, I6, I8–I9, I11, I18) to jointly develop a solution incrementally. One interviewee explained: “When structuring, we use [GenAI] as double check in the sense of what does [the GenAI] suggest, or what can be optimized further” (I11). Many interviewees also use GenAI at the end of the task, such as to check grammar (e.g., I8, I12–I13), to help writing a text in a different language (e.g., I6, I9, I17), or to perform last sanity or completeness checks on a work product like a presentation (e.g., I11, I21).

*Situational GenAI utilization.* Maintaining cognitive resources is a core idea of smart loafing (Stieglitz et al. 2022). Interviewees (e.g., I5, I17) specifically reported

using GenAI when feeling low in energy, lacking the mental capacity to tackle certain tasks, particularly in the hour before lunch and in the afternoon hours. They then used GenAI to create a first draft of the work product and return to it in times of higher energy. Starting to work on a new topic at the border or outside of one's own area of expertise can be time-consuming. Several interviewees have faced this issue and used GenAI to ease their start into the new topic when seeking to build up expertise (e.g., I5, I9, I15, I18–I19). One participant (I18) explained to that end: “Sometimes, when I'm under time pressure and I have a new topic where I don't yet have a complete overview or perhaps, I want to get into the technical details quickly and easily, I use ChatGPT as a bit of a sparring partner to get a better overview of the topic”. Another participant (I20) offered more nuance by saying that their use of GenAI depends on how sophisticated and complex the task is, the more sophisticated the task, the less likely they were to use GenAI.

If they need to perform a onetime task in which they are not an expert, they might also use GenAI so they do not have to build up the skills they would need to complete the task by themselves (I6, I10), e.g., to break down complicated, domain-specific KPIs or to apply legal regulations.

It becomes apparent that participants deliberately choose how, when, and why to engage GenAI, with consequences to effort intensity (e.g., trying less hard for tasks which can be covered by GenAI), effort persistence (e.g., overcoming the blank page or countering overthinking), and effort direction.

## 5.6 Work Effectiveness and Efficiency

Effort and performance are correlated (van Iddekinge et al. 2023), and according to *smart loafing*, enhancing work efficiency by relying on the technical system is a key concern (Stieglitz et al. 2022).

*Time efficiency (↑)* Interviewees reported improvements in both effectiveness (tackling harder tasks, I9) and efficiency (e.g., I1, I3–I4, I6, I9, I12–I15, I19–I21), with the emphasis on faster task completion. Reported subjective efficiency gains ranged between 60 and 90% in some cases (e.g., I6, I14) with one participant noting: “If I didn't have ChatGPT, I'm not sure if I would have actually managed the workload that I'm managing now” (e.g., I14). Interviewees realized these gains by GenAI helping them to prevent them from reading irrelevant text, speeding up routine work (e.g., I4, I14–I15), researching topics difficult to google (I9, I12, I19), or composing documents by copying large amounts of GenAI output. However, efficiency gains seem to be much smaller for more conceptual tasks, one interviewee estimated around 3–5% improvement (I5), as much of the text is too generic and requires

dedicated research and enrichment with expert knowledge. Some interviewees also drew comparisons with humans (e.g., working students), explaining that using GenAI was probably more cost-efficient (I10), the communication less ambiguous and the work results more neutral (I6).

*Quality (?)* The interviewees have mixed perceptions on how quality is affected by using GenAI. Some reported instances of quality improvement both in form and content, e.g., GenAI phrasing contents more eloquently (e.g., I9, I17), or cases in which GenAI proposed additions to their (conceptual) work results they had not thought about, but which were received positively by them or others (e.g., I6, I11). Others mentioned that the quality expectations and consequently the quality remained unchanged. Some interviewees felt that quality was slightly lower compared to working alone (e.g., I2) for certain tasks (e.g., I15). However, one interviewee (I2) mentioned the Pareto principle explaining that for many work settings, even 80% quality solutions are sufficient, and particularly given the saved time, a reduction in quality can be acceptable.

## 5.7 Cognitive Resources

According to smart loafing, a key idea of relying on technical systems is to maintain cognitive resources when completing a task (Stieglitz et al. 2022). While not frequently mentioned explicitly as a reasoning for using GenAI, even when asked about it directly, interviewees shared some reflections related to cognitive resources.

*Relief through GenAI (↓)* Most interviewees felt using GenAI to be a relief. They used it, when they had insufficient mental capacity to work on a task (e.g., I5, I19) or they felt lazy (e.g., I17). It helped freeing them from routine or operational work, providing them with more time to perform other tasks like strategic, creative or conceptual work (e.g., I9, I11–I13). Most participants explained that, if they did not perceive using the system as a relief, they would not use it and that doing the same work alone (as compared with GenAI) would require more “brain work” (I10, I19, I20), indicating at a potential reduction of cognitive load.

*Stable external expectations (→)* However, participants explained that the fact that they use GenAI does not change the stakeholders’ expectations, or, as an interviewee (I2) put it: “The pressure comes more of the deadline or stakeholders and they don’t really care how the product came about [...], so they don’t care whether I spent 5 h on it or 100 h, the quality has to be right. And that goes hand in hand with my feeling that my pressure or stress [...] hasn’t really changed. Just because you have a model like

this, you can work faster, but the expectations for the result remain the same.” While GenAI can support task aspects, it also produces output requiring verification, creating additional cognitive load (e.g., I17, I19). Because result expectations still must be met, and – as before – the worker is held accountable for the results, the cognitive load or stress remains the same (e.g., I2, I6), irrespective of how the results are produced. This point is closely connected to the external responsibility mentioned in the section on perceived responsibility.

*Continuous learning (↑)* Several participants reported putting in additional effort to learn to use the tools effectively (also see section on expectancy). Some interviewees mentioned aspects which do or could contribute to an increase in cognitive load, because they not only have to learn to use GenAI in the beginning but to continuously learn how to use it (e.g., I11, I17). Additionally, they have to test new versions, models, or modules to assess their potential for certain (business) tasks and business functions (e.g., I5, I9–I10, I16), which some described as exhausting (e.g., I5, I16). Furthermore, several interviewees shared that GenAI’s output is not always satisfactory to them. While not explicitly described as a burden, multiple interviewees report that when the output of GenAI is not satisfactory, they believe that they have made a mistake (e.g., I9, I11, I8) and that the system is “cumbersome” (e.g., I14). This continuous learning can be perceived as additional (cognitive) effort, besides working on the task. One participant (I4) how the continuous learning affected her as follows “More and more [tools] are being thrown onto the market and once a week, someone comes around the corner with a different idea that you could try out [...] I wonder if someone else can try it out and tell me how well it works, because I find it exhausting when you try to understand a new tool every week and have to test out where the limitations are.”

*Additional task (?)* Lastly, different layers need to be considered regarding relieving professionals through GenAI: task- vs. job-level. Many interviewees state that they feel relieved and save time due to using GenAI for the tasks they perform. However, when asked about what they do with this saved time, they usually do not invest it into enhancing the quality, nor do they use it as free time, but rather they use it to tackle additional new tasks, or as one interviewee put it: “I have so much on my plate [...] there’s always something else on the To-Do list filling up the 30 min of saved time” (I12). Thus, while relief might occur on the individual task level, on the overall job level, there might be a “fill-up” effect. Only two interviewees

reported that using GenAI gives them some time off in between tasks or for having a coffee (I10–I11).

## 6 Discussion

### 6.1 GenAI's Effect on Human Behavior and Relationships

Our findings reveal three linked insights into how working with GenAI can change human behavior, specifically, how GenAI can reshape human motivation (1) and how humans invest effort (2). Additionally, the results indicate that human relationships in work contexts are related to working with GenAI (3).

*Reshaped motivational antecedents* Expectancy theory (Vroom 1964) and the Collective Effort Model (Karau and Williams 1993) predict that when workers believe that their effort will not meaningfully enhance performance, e.g., when working with a highly capable partner, workers will withhold effort. Prior studies extended this logic to AI predicting effort reductions (Liu et al 2023), akin to social loafing in human groups, concluding that “[social loafing] occurs in virtual collaboration with [virtual assistance]” (Stieglitz et al 2022, p. 759).

In our study, participants working with GenAI feel their effort to be dispensable for some (sub-)tasks. Overall, however, they report that they can still improve performance either by working on the work product directly or indirectly through prompting (*expectancy*), sometimes even feeling enabled. They perceive their involvement to be essential to achieve desired outcomes (*instrumentality*), stressing their process administration role as well as the necessity to provide context-specific knowledge and to shape and communicate the work product to meet stakeholder expectations. This contrasts with expectations of eroded necessity (Liu et al 2023).

Social loafing in human groups frequently stems from diffused responsibility, when contributions are integrated into a single work result and are not identifiable. Earlier work suggests that such responsibility diffusion may not be limited to human groups (Henkenjohann and Trenz 2024; Stieglitz et al. 2022), but that humans may similarly assign responsibility to AI, stating that “the responsibility of solving a task in virtual collaboration is likely to be attributed to a [virtual assistant]” (Stieglitz et al. 2022, p. 759). Stieglitz et al. (2022, p. 759) concluded that for “virtual collaboration with [virtual assistants]” their study “indicates that knowledge from human-to-human collaboration in terms of [social loafing] is transferable”. Our data, however, suggests that social loafing-related knowledge from human groups is not necessarily transferable to professional human-GenAI settings when it comes to the

identifiability of contributions and diffused responsibility. Potential explanations might be found in the specific characteristics of GenAI and its contributions, as well as in the fact that the tasks in our study are embedded in real work contexts in relation to other tasks, fellow humans and organizational expectations.

Literature on social loafing shows that when individual inputs merge into a single group product and cannot be identified, *responsibility diffuses* and effort tends to be reduced (i.e., social loafing occurs, Karau and Williams 1993; Latané et al. 1979). If, on the other hand, an individual's contribution is (a) *identifiable* or *known*, and (b) there is a *standard* against which the contribution can be compared, such loafing behavior may be counteracted (Karau and Williams 1993). In our study, many interviewees report that recipients may struggle to separate their contribution from the GenAI's (i.e., contributions are not *identifiable*) as GenAI outputs often resemble human contributions. Due to the iterative nature of working with GenAI, with constant re-integration of own and genAI contributions, individual contributions might be even more difficult to discern, as was also suggested in prior literature on human-AI teaming (Rix and Hess 2022). Additionally, the interviewees do not always disclose having used GenAI (i.e., contributions are not *known*). According to classical social loafing literature, this might lead to diffused responsibility. However, most interviewees report to still feel fully responsible. They cite both a professional ethos and explicit organizational expectations (i.e., a *standard*): the final work result carries their name, regardless of how much help the GenAI provided. This finding is contrary to prior studies, which showed that humans assigned responsibility to AI (Henkenjohann and Trenz 2024; Stieglitz et al. 2022).

Potential explanations for the divergence of our findings compared to predictions and findings on human-AI dyads (Henkenjohann and Trenz 2024; Liu et al. 2023; Stieglitz et al. 2022) may be the contextual factors. First, the specifics of GenAI differ from the technical systems used in prior studies. Producing high-quality output requires iterative prompting: users must steer, enrich, and repeatedly verify responses that are probabilistic and potentially hallucinated. This sustained, hands-on work and the need to integrate results into the organization keeps human's expectancy and instrumentality high. Second, unlike brief lab tasks, our participants operate in authentic professional settings, where deliverables bear their name, and serve real stakeholders; established accountability norms anchor responsibility to humans, counteracting diffusion observed in controlled experiments (Henkenjohann and Trenz 2024; Stieglitz et al. 2022; also see section “GenAI usage and relationships with others”).

For now, GenAI appears to reshape rather than erode motivation with antecedents of effort remaining largely intact, even as workers incur new learning and verification burdens (see next section). Figure 2 synthesizes qualitative insights and provides a basis for testable propositions such as (exemplary selection):

**Proposition 1** *The stronger a worker’s belief that learning to steer GenAI improves its output, the higher their expectancy and the lower their intention to reduce effort for the task.*

**Proposition 2** *The stronger a worker’s belief that human process administration is indispensable when working with GenAI, the higher their instrumentality and the lower their effort-reduction intention for the task.*

**Proposition 3** *If workers expect an unshared reward (i.e., full personal credit) when working with GenAI, the perceived valence (i.e., the desirability of the outcome) remains as high as for unaided work and they do not reduce their effort on the task.*

**Proposition 4** *The stronger an organization’s accountability norms, the less probable it is that work with GenAI will be felt to reduce responsibility.*

Suggested antecedents and relationships in Fig. 2 also highlight conditions under which effort reduction could emerge, e.g., if future systems (a) consistently outperform humans even in expert tasks – lowering *expectancy*; (b) autonomously orchestrate workflows and interface with stakeholders – lowering *instrumentality*; (c) lead colleagues to withhold credit for outputs that can be generated by GenAI – lowering *valence*; or (d) are perceived as genuine teammates (Seeber et al. 2020), enabling *responsibility* diffusion. Tracking such technological and social shifts will be essential for anticipating when effort investment might slide toward loafing-like behavior.

**Effort investment beyond effort reduction.** Prior literature on human–AI work has used a social-loafing lens (Liu et al. 2023; Stieglitz et al. 2022), predicting – or experimentally showing – lower human effort when working with AI, and has recently refined that idea into smart loafing: a deliberate effort reduction to gain efficiency and conserve cognitive resources (Stieglitz et al. 2022). Our interviews partially confirm that pattern on a (sub)task level: knowledge workers often reduce effort intensity for (sub-)tasks which GenAI handles well. However, participants also report that GenAI helps them to overcome hurdles in longer work processes – such as low energy or the “blank page” problem – thus maintaining momentum in more cognitively demanding tasks and increasing their *persistence*. We therefore propose:

**Proposition 5** *Working with GenAI increases effort persistence.*

Perhaps more importantly, many report to deliberately shifting their effort *direction* by selectively using GenAI for certain (sub-)tasks (e.g., initial draft generation) and investing the freed-up time in higher-level thinking for the same or an additional task; hence, we propose:

**Proposition 6** *When working with GenAI in a job context, reduced effort intensity for certain sub-tasks is – at least partially – offset by redirected effort into other sub-tasks or additional tasks.*

This seems to result in a “fill-up” effect on job-level, where time saved on one task is used to perform additional tasks. Although it is unclear whether working on these additional tasks adds (similar) value as compared to working more intensively on the original task, this behavior is a deliberate, context-dependent approach to allocating effort, rather than a simple, across-the-board reduction in effort as the term “loafing” might suggest. Extant literature has investigated effort reduction behavior in experiments in which participants had to perform one task several times in a row (e.g., Stieglitz et al. 2022); future studies should consider investigating settings containing *multiple, different* tasks to uncover the dynamic effort redirection in GenAI work (cf. Dell’Acqua et al. 2023).

A further layer of effort emerges when considering the uncertainty GenAI introduces to the effort-outcome link. When people work alone, the work result is largely dependent on their effort; with GenAI, it depends on two additional contingencies. First, the GenAI must – in principle – be capable of producing the content for the desired outcome. Second, the user must be able to elicit that content through effective prompting. To reduce uncertainty, respondents invest effort up front and continuously: they build prompt libraries, read blog posts, follow tutorials, and iteratively work with GenAI. These activities often pay off but are sometimes a time sink – time lost to hallucinations, style mismatches, or fruitless prompt tweaks. This is consistent with findings from Dell’Acqua et al. (2023), where a dedicated GenAI training (contrary to expectations) led to a performance reduction. This parallels the contingencies highlighted in the Collective Effort Model (Karau and Williams 1993): just as performance in human groups depends on others, performance in human–GenAI dyads depends on an uncertain technical “partner”. A purely outcome-based perspective risks failing to acknowledge the specifics of working with quickly evolving, probabilistic GenAI systems, in which (up-front and continuously) invested effort does not guarantee improved performance (i.e., additional effort does not necessarily yield better results). When workers let GenAI draft texts,



the visible drop in keystrokes may coincide with a redirection of effort to strategic analysis (i.e., intensity reduction at task level, but effort redirection at job level).

Overall, we argue for shifting the focus from asking whether humans reduce effort to examining how and why they invest it. A broader lens therefore needs to be multi-dimensional (different effort dimensions), multi-level (task and job), and longitudinal (capturing up-front learning and ongoing prompt-tuning). It should acknowledge the contingencies introduced by GenAI, where more effort does not always yield better results. Calling such behavior loafing – defined by Merriam-Webster (2024) as “to spend time in idleness” – risks mis-classifying effort re-allocation as idleness. A vocabulary centered on effort investment might better capture knowledge workers’ GenAI usage behavior.

*GenAI use and relationship with others* A difference of our compared to prior experimental studies is that our study is embedded in work contexts with real relationships to the recipients of the results produced with GenAI. Though not at the core of our study, our data suggests that human relationships might influence and be influenced by the work with GenAI beyond the impact of GenAI on the individual human (behavior). For example, this could occur with regard to *contribution identifiability* and *GenAI-use disclosure*, *unshared rewards*, or *satisfactory GenAI output*.

As discussed in the section on identifiability, despite human and GenAI contributions being closely integrated, workers still feel responsible for these work results due to organizational norms and stakeholder expectations. Hence, the decision to disclose or not disclose GenAI usage seems tactical rather than a way to enable social loafing. Interviewees decide on disclosure or non-disclosure based on factors such as the relationship with the recipient, or the importance or stage of the work result. Disclosure may help to manage expectations or spread good practice; non-disclosure may feel appropriate when GenAI’s role was minimal (akin to a Google search). In short, selective disclosure is tied to a variety of contextual motives and is not necessarily a means of evading effort. This relational pattern differs from human-only groups: even with low identifiability, responsibility can remain firmly anchored to the human so long as organizational norms or stakeholder relationships assign it there.

Relational dynamics extend to rewards and refining GenAI outputs. Interactions with colleagues influence motivational drivers (and consequently effort), e.g., if the praise or *reward* a human receives is made dependent upon whether and to which degree a work result was produced with a GenAI, this may affect the *valence* (i.e., the desirability of the outcome) of the individual sharing the result. Additionally, stakeholders’ quality expectations may affect

whether GenAI output is considered *satisfactory* or needs further *refinement*, affecting the person working with GenAI. Inversely, when sharing work results without sufficient quality assurance, workers may transfer workload to others or introduce risks. These factors have important implications for managing GenAI usage in organizations. Future research may employ organizational case studies – examining sender–recipient pairs or networked interactions – to reveal how social context mediates GenAI use and effort allocation.

## 6.2 Contribution to Theory

First, our study details how motivational antecedents are reshaped when knowledge workers use GenAI. Laboratory studies have predicted (Liu et al. 2023) or demonstrated effort reduction under certain controlled conditions and concluded that social loafing “occurs” when working with AI (Stieglitz et al. 2022, p. 759). Our study nuances this claim for a professional work context, finding that expectancy, instrumentality, and responsibility can stay high, while valence can remain stable, so a behavior that resembles loafing is not inevitable. Figure 2 offers a detailed view on how GenAI can increase, decrease, or redirect these antecedents based on its capabilities and usage. The model offers a testable map for future experiments and can serve as an aid for managers seeking to understand when and how GenAI is likely to erode or amplify motivation.

Focusing on a specific antecedent of effort reduction – diffused responsibility – our findings (1) reveal a potential boundary condition that complicates the classic link between identifiability and perceived responsibility, and (2) challenge prior literature that observed a shift of responsibility toward AI. In human groups, low identifiability of inputs allows responsibility to diffuse, setting the stage for loafing (Karau and Williams 1993). When working with GenAI, iterative prompting can weave GenAI and human contributions into a single result, i.e., the human contribution may have low identifiability – yet professionals continue to feel responsible. We suggest that this finding which is contrary to prior studies (cf. Henkenjohann and Trenz 2024; Stieglitz et al. 2022) may be due to organizational norms: low identifiability does not necessarily undermine responsibility if the “other member” is a GenAI and the organizational context keeps credit and blame with the individual. Future research should treat identifiability and transfer of responsibility as contingent on local accountability norms; only when norms allow diffusion should indistinguishability translate into lower responsibility and effort. Selective disclosure of GenAI use therefore does not imply an evaded effort, but it occurs as a situational signal.



Secondly, we reframe the recent discourse beyond effort reduction on the level of a specific task, towards a (broader) *effort investment lens*. Prior, lab-based studies measured effort (or loafing-like behavior) on a focal task and linked reduced effort to saved cognitive resources (Stieglitz et al. 2022). Our data confirms such reduced effort on a (sub-)task level but shows that workers may redirect time to – according to their perspective – higher-value (sub-)tasks or additional tasks, potentially resulting in a job-level fill-up effect. Moreover, participants invest considerable effort into improving their skills. Understanding such patterns requires tracking intensity, persistence, and direction of effort across tasks and over time, and factoring in the probabilistic, evolving nature of GenAI. Rather than abandoning loafing theory, we therefore suggest a more comprehensive *effort investment* perspective, shifting the core question from “Does effort become less?” to “Where is effort invested, and why?” – a question that calls for longitudinal, multi-dimensional analyses able to distinguish idleness from strategic re-allocation. While controlled experimental studies are important to measure effort on specific tasks, our qualitative study sheds light on the broader interrelations in work settings. This study provides indication that an observed reduction of effort on a focal task might have different reasons, including not only behavior analogous to loafing, but also investment in future skills or shifts to other tasks. Additionally, the terminology of *investment* reflects the common belief among our participants that the effort spent to learn how to effectively work with GenAI will pay off, while acknowledging the uncertainty of returns related to probabilistic, quickly evolving GenAI systems.

### 6.3 Implications for Practice

Our findings highlight that working with GenAI in professional settings can reshape – rather than merely reduce – human motivation and effort. However, whether workers can leverage GenAI’s benefits depends on contextual factors: policies, expectations, and relationships must adapt to reflect the specifics of working with GenAI. These specifics include the quickly evolving capabilities of GenAI and the uncertainty surrounding working with GenAI, where more effort does not always yield better results. Managers play a critical role in shaping these conditions, e.g., by designing supportive environments, clarifying responsibilities, and enabling effective use of GenAI.

*Encourage exploration, but acknowledge learning costs and uncertainty* Aligning with our broader conceptualization of effort investment, managers should recognize that up-front and continuous learning in effectively steering GenAI and integrating it into the workflow is critical for leveraging GenAI effectively. Encouraging employees to

experiment and allowing for dips in efficiency may pay off. Viewing inefficiencies as “wasted effort” does not acknowledge the contingencies associated with using ever changing, probabilistic GenAI for knowledge work, where additional effort may not always translate into improved performance. A supportive environment – where employees feel safe to iterate and even fail – can ultimately support the development of competencies for GenAI usage.

*Clarify expectations and responsibilities* Organizations seeking to develop policies for GenAI usage should carefully consider the new ways of working with GenAI. Knowledge workers adjust the effort they invest into tasks, how much they use GenAI, and how much effort they spend on verifying GenAI outputs dynamically, enabling them to quickly iterate on their path toward final work results. Static policies like “You are responsible for the quality of your results” may not adequately reflect the dynamic use of GenAI to quickly produce drafts results to share with colleagues that we have observed in our data; instead, it may prevent workers from using GenAI if they do not have the expertise or time to verify outputs. A more flexible approach, e.g., asking workers to communicate the stage or quality of the shared result, may be more helpful, as it allows workers to benefit from the potential of GenAI while helping to prevent potential risks or causing additional work.

*Situational disclosure* Our findings suggest that decisions around disclosing GenAI usage are rarely about evading effort; instead, they serve as contextual signals that help to communicate quality levels, manage expectations, or advocate GenAI use within teams. Therefore, mandating universal disclosure of GenAI use may be counterproductive. It could inadvertently lower motivation (valence) if colleagues or managers assume that “the GenAI did all the work,” diminishing perceived individual contribution. It might also encourage “lazy draft” behavior where workers offload low-quality outputs onto others, assuming disclosure shifts the burden of review. Rather than blanket policies, organizations should consider implementing flexible, context-sensitive disclosure guidelines. These should support building a shared understanding of a work result’s quality, allowing workers to benefit from GenAI while maintaining responsibility in collaborative settings.

### 6.4 Limitations and Further Research

This study employed a qualitative, interview-based approach, which poses certain challenges in assessing changes in effort. First, social desirability bias (Döring and Bortz 2016) may have led participants to only mention selected interactions or downplay reductions of effort, especially given the negative connotations of “loafing.” To mitigate this, we promised anonymity to the interviewees,

built rapport prior to addressing sensitive topics, explained that we seek to understand their experience, and asked indirect questions, e.g., also inquiring about antecedents and considering behaviors, even if participants did not label them as “reduced effort” directly. Second, participants might unconsciously reduce effort, truthfully reporting their experiences but remaining unaware of underlying effort shifts. We tried to counter this and to uncover effort changes that interviewees had not been aware of during the situation (i.e., unconscious effort changes) by asking them to walk us through specific tasks completed with GenAI and reflecting on their behavior to better understand their decision-making and potential “preexisting effort scripts” (Karau and Williams 1993, p. 685). Nonetheless, potential unconscious effort changes may have been missed by our approach; future research might mitigate this through triangulation, e.g., using trace data or in-situ observations. Third, to reduce subjectivity during the analysis, we followed a directed content analysis approach with predefined categories (based on our conceptual model) and incorporated representative quotes; nonetheless, potential biases persist. Lastly, for our directed content analysis, we derived a conceptual model integrating previously discussed antecedents and consequences for effort investment when working with AI. However, literature on expectancy theory and social loafing is vast and we encourage exploring additional factors.

We conducted interviews with a relatively small number of knowledge workers spanning various professional domains, though nearly half were researchers. While we reached saturation – no new insights emerged from subsequent interviews – the overrepresentation of researchers may have caused us to conclude saturation prematurely, and interviewing additional participants in other roles or industries might uncover different or more nuanced behaviors. Moreover, most interviewees primarily used ChatGPT or Microsoft Copilot, which could influence our findings. However, as these systems are widely used, they might be representative of professionals’ experiences. While participants varied widely in their prevalent tasks, we drew on established theories (expectancy theory, social loafing) that have been validated across diverse settings. The broad sampling with explicit selection criteria determined by the study goal aimed to capture general effort-related patterns rather than confining the study to one domain.

Our data reflects participants’ perceptions at specific points in time, during which organizations and individuals are still adapting to GenAI. Future, more capable or agentic systems which are able to autonomously discover tasks and interact with stakeholders could further alter motivation or reduce perceived responsibility, making loafing-like behavior more likely. Additionally, changes in perception

of GenAI-assisted work may shift valence or instrumentality. Our model (Fig. 2) may serve as a starting point to (quantitatively) study the effects of these developments on human motivation and effort.

Moving forward and building on the “smart loafing” concept (Stieglitz et al. 2022), future research should consider *effort investment* when working with GenAI, also reflecting the additional contingencies that occur. Such work might examine longitudinal changes in effort distribution as people gain proficiency, capture real-time GenAI usage data, or investigate organizational practices (e.g., disclosure policies) that shape how humans balance their tasks with GenAI support. Additionally, investigating sociotechnical solutions for sharing GenAI-assisted outputs – such as standardized labeling for draft quality or embedded revision histories – could prevent miscommunications, manage stakeholder expectations, and mitigate potential additional effort or risks. Embracing these avenues will allow future IS studies to deepen our understanding of how knowledge workers meaningfully integrate rapidly evolving GenAI into their work, thereby preventing some of the aforementioned challenges.

## 7 Conclusion

GenAI is increasingly woven into professional knowledge work and contrary to fears of widespread effort reduction, our study suggest that knowledge workers deliberately manage their effort: while they rely on GenAI for certain subtasks, they also invest effort in learning to prompt, verify, and integrate GenAI outputs. Simply transferring the social loafing lens from purely human groups to human–GenAI dyads risks overlooking this dynamic interplay – where up-front learning, strategic reallocation of effort, and an uncertain relationship between worker input and GenAI-generated outputs all come into play. We thus propose a more nuanced perspective, highlighting ways of how GenAI might reshape effort investment both directly and indirectly, and offer initial practical implications for managers and organizations. However, as GenAI capabilities evolve, further research is needed to refine our understanding of how professionals can optimize their effort investment – maintaining responsibility, enhancing performance, and upholding motivation – in an era of rapidly advancing GenAI.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12599-025-00960-4>.

**Acknowledgements** This research was funded by the German Federal Ministry of Education and Research (BMBF) in the context of the project HyMeKI (reference number: 01IS20057).

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Cymek DH (2018) Redundant automation monitoring: four eyes don't see more than two, if everyone turns a blind eye. *Hum Fact* 60(7):902–921
- Cymek DH, Manzey D (2022) Sequential human redundancy: Can social loafing diminish the safety of double checks? *J Exp Psychol Appl* 28(4):931–945
- Cymek DH, Trukenbrodt A, Onnasch L (2023) Lean back or lean in? Exploring social loafing in human-robot teams. *Front Robot AI* 10:1249252
- Dell'Acqua F, McFowland E, Mollick ER, Lifshitz-Assaf H, Kellogg K, Rajendran S, Kraye L, Candelon F, Lakhani KR (2023) Navigating the jagged technological frontier: field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Working Paper*
- Dellermann D, Ebel P, Söllner M, Leimeister JM (2019) Hybrid intelligence. *Bus Inf Syst Eng* 61(5):637–643
- Döring N, Bortz J (2016) *Forschungsmethoden und evaluation in den Sozial- und Humanwissenschaften*. Springer, Heidelberg. <https://doi.org/10.1007/978-3-642-41089-5>
- Guest G, Bunce A, Johnson L (2006) How many interviews are enough? An experiment with data saturation and variability. *Field Meth* 18(1):59–82
- Haase J, Hanel PHP (2023) Artificial muses: generative artificial intelligence chatbots have risen to human-level creativity. *J Creat* 33(3):100066
- Hamza M, Siemon D, Akbar MA, Rahman T (2024) Human-AI collaboration in software engineering: lessons learned from a hands-on workshop. In: 7th ACM/IEEE International Workshop on Software-intensive Business. ACM, pp 7–14
- Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: *Advances in Psychology*. Hum Ment Workload 52:139–183
- Henkenjohann R, Trenz M (2024) Challenges in collaboration with Generative AI: interaction patterns, outcome quality and perceived responsibility. In: *ECIS 2024 Proceedings*
- Hsieh H-F, Shannon SE (2005) Three approaches to qualitative content analysis. *Qual Health Res* 15(9):1277–1288
- Hu K (2023) ChatGPT sets record for fastest-growing user base. Reuters. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
- Hüffmeier J, Dietrich H, Hertel G (2013) Effort intentions in teams. *Small Group Res* 44(1):62–88
- Järvenpää E, Eloranta E (2001) Information and communication technologies and quality of working life: Implications for competencies and well-being. In: Bradley G (ed) *Humans of the net – Information & communication technology (ICT), work organisation and human beings*. Prevent, Stockholm, pp 109–118
- Karau SJ, Williams KD (1993) Social loafing: a meta-analytic review and theoretical integration. *J Pers Soc Psychol* 65(4):681–706
- Latané B, Williams K, Harkins S (1979) Many hands make light the work: the causes and consequences of social loafing. *J Pers Soc Psychol* 37(6):822–832
- Liu M, Ke W, Xu DJ (2023) Will humans be free-riders? The effects of expectations for AI on human-AI team performance. In: *PACIS 2023 Proceedings*. <https://aisel.aisnet.org/pacis2023/20>
- Makarius EE, Mukherjee D, Fox JD, Fox AK (2020) Rising with the machines: a sociotechnical framework for bringing artificial intelligence into the organization. *J Bus Res* 120:262–273
- Mandt E (1978) Managing the knowledge worker of the future. *Pers J* 57(3):138–162
- Mayring P (2014) *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Klagensfurt. <https://nbn-resolving.org/urn:nbn:de:0168-ssao-395173>. Accessed 10 Oct 2025
- Memmert L, Tavanapour N, Bittner E (2023) Learning by doing: educators' perspective on an illustrative tool for AI-generated scaffolding for students in conceptualizing design science research studies. *J Inf Syst Educ* 34(3):279–292
- Memmert L, Bittner E (2022) Complex problem solving through human-AI collaboration: literature review on research contexts. In: 55th Hawaii International Conference on System Sciences. [https://aisel.aisnet.org/hicss-55/cl/machines\\_as\\_tammates/3/](https://aisel.aisnet.org/hicss-55/cl/machines_as_tammates/3/)
- Memmert L, Bittner E (2024) Human-AI collaboration for brainstorming: effect of the presence of AI ideas on breadth of exploration. In: *HICSS57 Proceedings*. [https://aisel.aisnet.org/hicss-57/cl/machines\\_as\\_tammates/6](https://aisel.aisnet.org/hicss-57/cl/machines_as_tammates/6)
- Memmert L, Tavanapour N (2023) Towards human-AI collaboration in brainstorming: empirical insights into the perception of working with a generative AI. In: 31st European conference on information systems. [https://aisel.aisnet.org/ecis2023\\_rp/429](https://aisel.aisnet.org/ecis2023_rp/429)
- Memmert L, Mies J, Bittner E (2024) Brainstorming with a generative language model: the role of creative ability and tool-support for brainstorming performance. In: *Proceedings of the 45th international conference on information systems (ICIS)*. <https://aisel.aisnet.org/icis2024/aiinbus/aiinbus/7/>
- Memmert L (2024) Brainstorming with a generative language model: understanding performance through brainstorming group effects. In: *ECIS 2024 Proceedings*. [https://aisel.aisnet.org/ecis2024/track06\\_humanaicollab/track06\\_humanaicollab/1](https://aisel.aisnet.org/ecis2024/track06_humanaicollab/track06_humanaicollab/1)
- Merriam-Webster (2024) Loaf. <https://www.merriam-webster.com/dictionary/loafing>
- Meuser M, Nagel U (2009) *Das Experteninterview – Konzeptionelle Grundlagen und Methodische Anlage*. In: Pickel S et al (eds) *Methoden der vergleichenden Politik- und Sozialwissenschaft*. VS Verlag für Sozialwissenschaften, pp 465–479
- Nass C, Moon Y (2000) Machines and mindlessness: social responses to computers. *J Soc Issues* 56(1):81–103
- Onnasch L, Panayotidis T (2020) Social Loafing with robots: an empirical investigation. *Proc Hum Fact Ergonom Soc Annu Meet* 64(1):97–101
- Pinsonneault A, Barki H, Gallupe RB, Hoppen N (1999) Electronic brainstorming: the illusion of productivity. *Inf Syst Res* 10(2):110–133
- Pyörä P (2005) The concept of knowledge work revisited. *J Knowl Manag* 9(3):116–127
- Retkowsky J, Hafermalz E, Huysman M (2024) Managing a ChatGPT-empowered workforce: understanding its affordances and side effects. *Bus Horiz* 67(5):511–523
- Rix J, Hess T (2022) Hello, mate! Insights from the field on leveraging machine teammates in organizations. In: *Pacific Asia*

- conference on information systems. <https://aisel.aisnet.org/pacis2022/95>
- Rix J (2022) From tools to teammates: Conceptualizing humans' perception of machines as teammates with a systematic literature review. In: 55th Hawaii international conference on system sciences. [https://aisel.aisnet.org/hicss-55/cl/machines\\_as\\_team\\_mates/5/](https://aisel.aisnet.org/hicss-55/cl/machines_as_team_mates/5/)
- Seeber I, Bittner E, Briggs RO, de Vreede T, de Vreede G-J, Elkins A, Maier R, Merz AB, Oeste-Reiß S, Randrup N, Schwabe G, Söllner M (2020) Machines as teammates: a research agenda on AI in team collaboration. *Inf Manag* 57(2):103174
- Siemon D, Eckardt L, Robra-Bissantz S (2015) Tracking down the negative group creativity effects with the help of an artificial intelligence-like support system. In: 48th HICSS proceedings. <https://doi.org/10.1109/HICSS.2015.37>
- Stieglitz S, Mirbabaie M, Möllmann NRJ, Rzyski J (2022) Collaborating with virtual assistants in organizations: analyzing social loafing tendencies and responsibility attribution. *Inf Syst Front J Res Innov* 24(3):745–770
- Sweller J (1988) Cognitive load during problem solving: effects on learning. *Cogn Sci* 12(2):257–285
- Timonen H, Paloheimo K-S (2008) The emergence and diffusion of the concept of knowledge work. *Electronic J Knowl Manag* 6:177–190
- Tsoukas H, Vladimirou E (2001) What is organizational knowledge? *J Manag Stud* 38(7):973–993
- van Iddekinge CH, Arnold JD, Aguinis H, Lang JWB, Lievens F (2023) Work effort: a conceptual and meta-analytic review. *J Manag* 49(1):125–157
- Vroom VH (1964) *Work and motivation*. Wiley, Hoboken
- Williams KD, Karau SJ (1991) Social loafing and social compensation: the effects of expectations of co-worker performance. *J Pers Soc Psychol* 61(4):570–581