The Role of Feedback in Cognitive Offloading during Human-Computer Collaboration

Jakob Trierweiler (j.trierweiler@campus.tu-berlin.de)
Institute of Psychology and Ergonomics, Berlin Institute of Technology, Berlin, Germany

Eva Wiese (eva.wiese@tu-berlin.de)
Institute of Psychology and Ergonomics, Berlin Institute of Technology, Berlin, Germany

Basil Wahn (basil.wahn@tu-berlin.de)
Institute of Psychology and Ergonomics, Berlin Institute of Technology, Berlin, Germany

Abstract

Cognitive offloading refers to the usage of physical actions to reduce the cognitive load of a task. This study investigates cognitive offloading in a multiple object tracking (MOT) task, in which participants could decide whether they wanted a computer partner to track targets on their behalf. Depending on the experimental condition, participants either received team performance feedback or not. In all conditions, participants knew that the computer partner would track targets flawlessly. We hypothesized that feedback would change participants' metacognitive assessments such that they would maximize performance (i.e., changing their strategy by offloading more or all targets to the computer partner). We find that participants do offload targets to the computer partner in all conditions. Yet, performance feedback did not increase the extent of cognitive offloading. While these findings dovetail with previous findings, future studies are needed to definitively rule out whether performance feedback has an effect on cognitive offloading.

Keywords: cognitive offloading; feedback; multiple object tracking; human-computer collaboration

Introduction

In today's digital age, humans are often faced with managing complex and cognitively demanding tasks. To handle these challenges, many turn to technological systems designed to support or even enhance their cognitive abilities. This interaction with technology can be understood through the lens of cognitive offloading, which refers to the usage of physical actions to reduce the cognitive load of a task (Risko & Gilbert, 2016). As Risko and Gilbert (2016) describe, this process can be seen as "Putting Cognition Into-The World". This aligns with foundational ideas in human-computer interaction research, particularly Don Norman's (Norman, 2013) principle of effectively combining "knowledge in the head with knowledge in the world" to optimize cognitive resources. Furthermore, theoretical perspectives on cognitive development and technological innovation suggest that human cognition operates in symbiotic interactions with digital agents (Clark & Chalmers, 1998; Clark & Erickson, 2004; Donald, 1993).

These frameworks help explain how external tools and systems can complement human cognition by distributing cognitive tasks between the individual and their digital environment. Cognitive offloading can be exemplified by someone using a calculator app for math or by someone using a calendar app on their smartphone to note down an appointment.

The integration of highly accurate technological systems into everyday tasks has raised important questions about how and when individuals choose to offload cognitive demands to them (Risko & Gilbert, 2016). The present behavioral study focuses on cognitive offloading onto technological systems. Previous research has already found factors that influence humans' cognitive offloading behavior. Key factors associated with increased cognitive offloading is the perceived accuracy (Wahn & Schmitz, 2024; Weis & Wiese, 2019, 2022) and task-fit (Hertz & Wiese, 2019; Wiese, Weis, Bigman, Kapsaskis, & Gray, 2022) of the digital agent with which individuals perform tasks. Another factor that influences cognitive offloading behavior is the presence of information about task-relevant proficiency of the digital agent. A prior study showed that participants allowed an app to solve a task for them more often if they were informed about task-related qualities of the app (Weis & Wiese, 2022). Additionally cognitive offloading behavior differs depending on the performance goals that are pursued. Weis and Wiese (2019) showed for an object rotation task, in which one must decide, whether two differently rotated stimuli are equal, that participants tend to rotate the stimuli internally while pursuing a speed performance goal while they rotate it more often with an external knob and a computer screen while pursuing an accuracy performance goal. Another factor that influences whether people engage in cognitive offloading is the presence of a bonus task. This was demonstrated in a prior experimental study (Wahn & Schmitz, 2024) where participants had to perform a multiple object tracking (MOT) task (Pylyshyn & Storm, 1988). The MOT task has an established role in exploring the limitations of attentional processing (Alvarez & Franconeri, 2007a). In the multiple object tracking (MOT) paradigm, participants are tasked with monitoring a select group of moving target objects amidst distractor objects displayed on a screen. Research has indicated that humans can typically track a finite number of objects with a high accuracy (Pylyshyn & Storm, 1988; Pylyshyn, 1989), often up to four (Intriligator & Cavanagh, 2001), though alternative findings suggest varying capacities depending on different parameters such as the speed of object motion (Alvarez & Franconeri, 2007b; Cavanagh & Alvarez, 2005; Feria, 2013). Tracking these objects is a demanding activity that needs continuous attention over extended periods (Alnæs et al., 2014; Wahn, Ferris, Hairston, & König, 2016). The MOT task has already been employed to study human-computer collaboration in various contexts, where participants could delegate some objects to a computer

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

partner, which would then track these objects on their behalf (Wahn & Kingstone, 2020; Wahn, Schmitz, Gerster, & Weiss, 2023; Wahn & Schmitz, 2024). Furthermore, performance in the MOT task has been shown to predict abilities in simulated air traffic control tasks (Jarvis, Hoggan, & Temby, 2022), highlighting its practical relevance and providing a foundation for the current study on cognitive offloading within this paradigm.

Previous results concerning cognitive offloading in the MOT task revealed that people engage in cognitive offloading, when possible, distributing mental workload between themselves and a digital agent to optimize cognitive resources and to achieve task-related goals (Wahn et al., 2023; Wahn & Schmitz, 2024), and participants' tracking accuracy increases when cognitive offloading is utilized, likely due to a reduction in individual cognitive effort (Wahn et al., 2023).

As technology – and particularly artificial intelligence – continues to outperform humans in an increasing number of tasks, understanding cognitive offloading behavior and the factors influencing it becomes crucial. This understanding is key to enhancing effective human-computer collaboration, especially for those responsible for developing and designing new digital solutions.

While performing tasks alone, humans can track their performance through self-monitoring mechanisms (Ullsperger, Fischer, Nigbur, & Endrass, 2014). Similarly, when tasks are performed in collaboration with either an external tool or another human, individuals can also monitor the performance of their collaboration partner or tool (Pfister, Weller, & Kunde, 2020; Weis & Kunde, 2024b). During human-computer collaboration, monitoring both one's own performance and that of the digital agent is crucial for deciding whether to offload cognitive workload. Prior research has highlighted the key role that assumptions about performance play in choosing a cognitive strategy (Gilbert, 2015; Weis & Kunde, 2024a). Thus, a major factor influencing cognitive offloading may be immediate feedback of the results after the end of a task, which could guide future cognitive strategy decisions when performing the task again, particularly regarding whether to engage in cognitive offloading or not. Feedback could inform individuals that their overall performance is better when they decide to delegate workload to a digital agent than attempting to perform a task alone. This information could guide individuals next strategic decisions resulting in a change in cognitive offloading behavior.

A prior study by Weis and Kunde (2024b) investigated the role of feedback using an object rotation task. Participants were asked to determine whether two presented stimuli were equal as quickly and accurately as possible. To reach their answer, participants could either mentally rotate (internally) the working stimulus presented on the right side of a computer screen or physically rotate it (externally) using a computer keyboard and visual support of the computer screen. They then compared it to the base stimulus presented on the left to determine equality or not. For most of the participants, there was not one optimal strategy that was faster and more accurate. However, Weis and Kunde (2024b) found that when participants received feedback either trialwise or after experimental blocks, they chose the more accurate strategy more often. However, this feedback did not lead to a significant increase in choosing the faster strategy.

Another study by Grinschgl, Meyerhoff, Schwan, and Papenmeier (2021) investigated whether fake-performance feedback in a memory-task influences cognitive offloading behavior. Its results indicate that metacognitive assumptions that are present in the strategy selection period during which individuals decide whether to engage in cognitive offloading or not (Risko & Gilbert, 2016) are influenced by fake performance feedback, however this impact did not lead to a change in cognitive offloading behavior (Grinschgl et al., 2021).

As of now, it remains unclear whether immediate feedback of the actual tasks' results influences future cognitive strategy choices, particularly when fully offloading the cognitive load would maximize the performance as seen in the MOT task with a 100% accurate computer partner (Wahn et al., 2023; Wahn & Schmitz, 2024). Immediate feedback could affect participants' metacognitive assessments, leading to more performance-maximizing behavior (i.e., offloading the entire task to the computer partner; Grinschgl et al. (2021); Risko and Gilbert (2016); Weis and Kunde (2024b)). Consequently, we hypothesize that feedback on the tasks' results will increase participants' cognitive offloading behavior during human-computer collaboration relative to a condition when performance feedback is not available.

The outcomes of this hypothesis have significant practical implications for the design of user interfaces, particularly in fields where collaboration with digital systems can enhance task performance. By understanding how feedback mechanisms influence offloading behavior, designers can create technological solutions that encourage users to rely more on digital agents when they outperform human capabilities. Providing feedback in certain human-computer environments could lessen individuals' technology aversion (Jussupow, Benbasat, & Heinzl, 2020; Wahn et al., 2023). This could lead to improved efficiency in environments where cognitive demands are high, such as air traffic control, medical diagnostics, or complex data analysis.

The hypothesis was tested with a behavioral experiment in which participants performed a multiple object tracking (MOT) task both individually and with a 100% accurate computer partner. Participants received feedback on their collaborative performance with the computer partner. The data from this experiment were then compared to data from our prior study, in which participants completed the same task without receiving performance feedback after each trial performed with the computer partner (Wahn et al., 2023).

Materials and methods

Participants

Twenty-six participants took part in the experiment (M = 33.85 years, SD = 14.44 years; 22 females (∼85%), 3 males (∼12%), 1 gender-diverse (∼4%)). An a priori power analysis was performed using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), determining a total sample size of 52 as sufficient for a repeated-measures ANOVA with one between-subjects and one within-subjects factor ($\alpha = .05, 1 - \beta = .94$) to detect medium-sized effects (Cohen's f = .25; Cohen (2013)). This total sample size of 52 for testing our hypothesis was achieved by combining data from the current experiment with data from our aforementioned study (Wahn et al. (2023); N = 26, M = 26.85 years, SD = 8.46 years; 18 females (∼69%), 8 males (∼31%) and no gender-diverse). All participants were recruited via the university's experimental participant pool and provided written informed consent before taking part in the study. For their participation they were compensated with experimental tokens, which are required for their degree in psychology. All participants agreed to the publication of their data.

Experimental setup and procedure

The same procedure as in our previous study (Wahn et al., 2023) was employed, with the only difference being the implementation of performance feedback provided after each trial of the MOT task when performed with a computer partner. Participants were seated at a desk, positioned 90 cm away from a 24-inch computer screen (1920 × 1080 pixels resolution, 30 frames per second). The keyboard and mouse were in proximity so that the participants could operate the computer.

Each session consisted of 75 trials of the MOT task, plus two training trials for each of the two experimental conditions (solo, joint). The experiment began with 25 trials in the solo condition and concluded with 50 trials in the joint condition. In the joint condition, participants performed the MOT task with a computer partner, allowing them to delegate targets to it that were then tracked on their behalf. The fixed order was chosen because, to make decisions about how much to offload, participants needed to be aware of their own individual capacity. Specifically, for the MOT task, they had to understand how many targets they could track on their own (Wahn et al., 2023). To facilitate this, participants first completed the solo condition, allowing them to assess their individual performance, followed by the joint condition, where they could then decide whether to reduce their tracking load compared to the solo condition by delegating targets to the computer partner.

During each trial (for a trial overview, see Figure 1) participants decided how many targets they wanted to attempt to track. After confirmation, targets were randomly assigned with a speed of 90 or 120 pixels per second. The targets moved for 11 seconds and subsequently the objects stopped moving. Participants were then instructed to click on the targets they had initially selected. The objective for participants was to accurately identify as many targets as possible while avoiding mistakes. Each correctly identified target earned one point, while each incorrect selection resulted in the loss of one point. Selecting more targets at the beginning of a trial for the tracking period than one was sure to safely identify in the target selection phase at the end of a trial did not lead to minus points. For example, if one attempted to track six targets in one trial but only was sure of the position of one of the six tracked targets after the tracking period, one point for a correct answer was scored. The trial sequence during the joint condition was similar to the one during the solo condition. Before performing the joint condition, participants were told that they are performing the next 50 trials with a computer partner which could track targets for them, if they decide to delegate cognitive load to it. Participants continued to choose 0-6 targets of the highlighted targets at the beginning of each trial but were informed that the computer partner would track the rest of the targets for them. They were told that their score consisted of the correctly identified targets by them (minus the incorrect selections) and by the computer partner and that the computer partner tracked targets 100% accurately. After delegating the workload for each joint trial, participants received visual feedback on the allocation of the targets. The targets that the computer partner would track during the trial were highlighted in pink. After the tracking period participants clicked on their tracked targets, marking them yellow and after confirming their "guesses" for the trial the computers' selection was displayed in pink. After pressing the space bar or clicking the dot in the middle of the screen the collaborative result (performance feedback) for the trial was shown. Participants mistakenly identifying a distracting target or a target allocated to the computer partner resulted in a minus point for the team score. And as for the solo condition participants did not receive minus points for attempting to track more targets than they could identify after the tracking period. The trial sequences of the solo and joint condition were the same as the ones in our previous study (Wahn et al., 2023) except for the introduction of the performance feedback of the team score, allowing the investigation of our hypothesis concerning the impact of the performance feedback on cognitive offloading.

As a point of note, given that using data from two different studies is not ideal, we want to point out that we ensured that experimental environments in which the present and our previous study (Wahn et al., 2023) were carried out are highly similar. That is, the data for both groups were collected using the same MOT task software in a comparable experimental environment (i.e., same monitor size and resolution), with both participant pools consisting of German university students.

Dependent variables

No data from specific trials or participants were excluded from the analyses. For each participant, mean values were calculated for the number of targets attempted to track during both the solo and joint conditions. These values were then

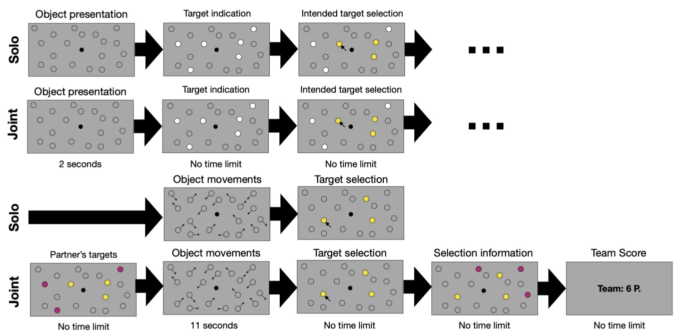

Figure 1: 1st and 3rd row: Depicts an example trial for the solo condition. 2nd and 4th row: Depicts an example trial for the joint condition.

averaged across all participants to obtain the sample mean for each condition. This process was repeated for the achieved points in each trial. Additionally, the accuracy for each trial was calculated by dividing the number of correctly identified targets by the number of targets participants attempted to track during that trial. The procedure was carried out separately for both the solo and joint conditions.

Results

To assess whether the data from the present and our previous study (Wahn et al., 2023) is comparable, we first compared the solo conditions with regard to the number of attempted targets, points scored, and accuracy across experiments. We reasoned that if experimental environments and samples are comparable between the present and previous study (Wahn et al., 2023), then participants should attain similar solo performances with regard to all three measures as the solo conditions did not change across studies. Indeed, we found performance to be similar for all three measures. That is, the number of attempted targets is similar and do not significantly differ ($t(50) = 1.36$, $p = .180$; With feedback: M = 3.32, SD = 0.92; Without feedback: M = 3.61, SD = 0.58). The same is true for the points scores ($t(50) = 1.29$, $p = .204$; With feedback: M = 2.68, SD = 0.50; Without feedback: M = 2.85, SD = 0.45) and the accuracy ($t(50) = 0.00$, $p = 1.00$; With feedback: M = 0.88, SD = 0.10 Without feedback: M = 0.88, SD = 0.06).

To assess whether participants offloaded targets in the joint condition relative to the solo condition and whether the offloading extent would differ across experiments (for a descriptive overview, see Fig. 2), we conducted a 2x2 ANOVA with the within-subjects factor Condition (solo, joint) and between-subjects factor Experiment (with feedback, without feedback). The dependent variable was the number of targets attempted to track in the beginning of a trial. Before performing the ANOVA, the assumptions of equal variances (Levene's test: $F(1, 102) = 0.03$, $p = .87$) and normally distributed residuals (Shapiro-Wilk test: W = 0.99, $p = .73$) were confirmed (Bortz & Schuster, 2010; Levene, 1960; Shapiro & Wilk, 1965). We found a main effect of Condition ($F(1,50) = 68.79$, $p < .001$), suggesting that participants offloaded targets in the joint condition. The main effect of Experiment ($F(1,50) = 1.06$, $p = .309$) and the interaction effect ($F(1,50) = 0.36$, $p = .550$) were not significant. Given that this interaction effect is central to our hypothesis, we assessed the evidence in favor of the null hypothesis. For this purpose, we computed the differences between conditions for each experiment and compared the differences using a Bayes factor. We found that the null hypothesis is 3.83 more likely for this comparison, which is considered substantial evidence in favor of the null hypothesis (Jeffreys, 1998), suggesting that the offloading extent does not differ when performance feedback is available or not.

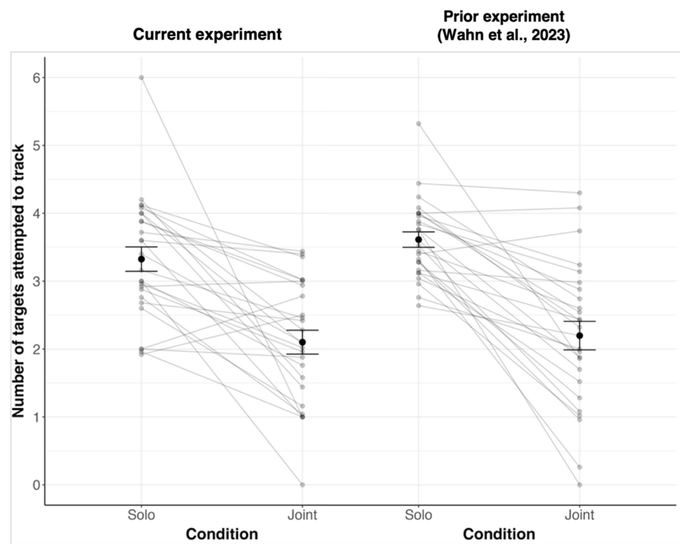

Figure 2: Average number of attempted targets to track for each condition (Solo vs. Joint), separately for the present study (with feedback; left panel) and a previous study (without feedback; right panel, Wahn et al. (2023); computer partner). Errors bars show Standard Error of the Mean. Lighter gray dots and lines are individual participants.

In addition, to test whether immediate feedback would affect the team's overall performance, we compared the team scores between experiments and found that they do not significantly differ ($t(50) = 0.82$, $p = .412$) and are close to the maximum possible performance, which is 6 (With feedback: M = 5.84, SD = 0.21; Without feedback: M = 5.78, SD = 0.31).

Discussion

The purpose of this study was to gain a better understanding of cognitive offloading during human-computer collaboration (Risko & Gilbert, 2016). We replicated that participants offload cognitive demand to a computer partner to perform a task more efficiently in line with our previous study (Wahn et al., 2023). The primary purpose of the study was to determine whether immediate feedback about the team's performance provided after collaborating with an accurate computer partner would enhance cognitive offloading behavior. This was done to better understand the conditions under which humans consult technology to achieve task-related goals. For the investigation of the effect that immediate performance feedback might have on cognitive offloading during human-computer collaboration, the current data were compared with that of our previous study (Wahn et al., 2023). The two experiments were completely equal except for the feedback of results provided after each trial that participants performed with the computer partner. Surprisingly, the results do not support the hypothesis and no significant effect of immediate performance feedback on cognitive offloading was found. Moreover, we found when calculating a Bayes factor, the null hypothesis is clearly favored. These findings suggest that, in our study design, immediate performance feedback does not influence cognitive offloading behavior onto a technological system.

To explain these results, it is possible that the performance feedback did not influence participant's metacognitive assumptions and assessments as well as the performance monitoring (Pfister et al., 2020; Ullsperger et al., 2014; Weis & Kunde, 2024b) and therefore no effect on strategy selection could be observed (Risko & Gilbert, 2016). This could mean that the performance feedback did not remind the participants of the performance maximizing strategy (offloading more or all targets to the computer partner). A future study could test whether a different form of communicating the performance feedback (i.e., highlighting whether the peak performance is reached or not) would be more effective to remind participants of the performance maximizing strategy.

Another potential explanation is that in the current study we did not create an environment in which the feedbacks' influence on metacognitive assumptions leads to a manifested change in strategy selection (more cognitive offloading behavior; Risko and Gilbert (2016)). Usually, individuals' strategic decision on whether to rely on internal cognitive resources or to delegate workload to an external tool is highly related to effort-performance tradeoffs (Cary & Carlson, 2001; Gray, Sims, Fu, & Schoelles, 2006; Grinschgl, Meyerhoff, & Papenmeier, 2020). However, the stakes in the current experiment were low and they could therefore affected participants' strategic decisions (Wahn et al., 2023; Wahn & Schmitz, 2024). In other words, participants felt no need to maximize the performance. This is supported by the fact that only few participants showed performance maximizing behavior by offloading almost every target to the computer partner and the other participants, potentially to avoid boredom and keeping cognitively engaged, attempted to track a handful of targets themselves (Wahn et al., 2023; Wahn & Schmitz, 2024). Wahn and Schmitz (2024) showed that when integrating an incentivised bonus task into the MOT paradigm therefore arguably raising the stakes, participants change their strategic decision towards a more performance-maximizing approach (more offloading). In short, in the current experiment participants did not follow a performance-maximizing approach because the results did not really matter to them. Therefore, they might have tended to pursue mastery goals rather than performance goals to potentially train their individual cognitive capacities (Dweck, 1986) or to stay engaged and to not feel bored (Wahn & Schmitz, 2024). Future research could further investigate the relationship between the perceived meaning of a situation and cognitive offloading behavior.

Another reason, why the results do not align with the predicted effect that immediate performance feedback might have on cognitive offloading behavior could be the difficulty of the MOT task in the current experimental design. Participants performed the task very accurately overall. So, from a performance-maximizing perspective it does not matter whether they track 0, 1, 2 or 3 targets themselves since they are close to 100% accuracy when attempting to track up to 3 targets (Intriligator & Cavanagh, 2001). A follow-up study could on the one hand raise the stakes and on the other hand the difficulty of the experiment, to create a quasi-real-life environment in which the potential effect of immediate performance feedback manifests a change in individual strategy selection (more cognitive offloading, Risko and Gilbert (2016)). One way to achieve a more difficult MOT task is to increase the speed of motion of all or some of the targets while e.g. highlighting the faster targets (Alvarez & Franconeri, 2007b; Cavanagh & Alvarez, 2005; Feria, 2013). In such a design the impact that immediate performance feedback has on individuals' metacognitive assumptions could manifest in a significant change in cognitive offloading behavior (Grinschgl et al., 2020; Risko & Gilbert, 2016).

In sum, the results of the current study provide clear support that performance feedback does not influence cognitive offloading behavior in an attentionally demanding task. These findings dovetail with recent findings in a different study Grinschgl et al. (2021), showing that performance feedback did not lead to a change in cognitive offloading behavior in a memory task. However, it is too early to rule out the potential of performance feedback for influencing cognitive offloading behavior and future studies could examine this question by increasing the stakes and/or task difficulty, or by changing how performance feedback is presented.

References

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of vision*, *14*(4), 1–1.

Alvarez, G. A., & Franconeri, S. L. (2007a). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13), 14.

Alvarez, G. A., & Franconeri, S. L. (2007b). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of vision*, *7*(13), 14–14.

Bortz, J., & Schuster, C. (2010). *Statistik für human-und sozialwissenschaftler 7th ed springer-lehrbuch.* Heidelberg: Springer Berlin) URL https://doi.org/10.17877/DE290R-6043.

Cary, M., & Carlson, R. A. (2001). Distributing working memory resources during problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 836.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*(7), 349–354.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19.

Clark, A., & Erickson, M. (2004). Natural-born cyborgs: Minds, technologies, and the future of human intelligence. *Canadian Journal of Sociology*, *29*(3), 471.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge.

Donald, M. (1993). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press.

Dweck, C. S. (1986). Motivational processes affecting learning. *American psychologist*, *41*(10), 1040.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Feria, C. S. (2013). Speed has an effect on multiple-object tracking independently of the number of close encounters between targets and distractors. *Attention, Perception, & Psychophysics*, *75*, 53–67.

Gilbert, S. J. (2015). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, *33*, 245–260.

Gray, W. D., Sims, C. R., Fu, W.-T., & Schoelles, M. J. (2006). The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. *Psychological review*, *113*(3), 461.

Grinschgl, S., Meyerhoff, H. S., & Papenmeier, F. (2020). Interface and interaction design: How mobile touch devices foster cognitive offloading. *Computers in Human Behavior*, *108*, 106317.

Grinschgl, S., Meyerhoff, H. S., Schwan, S., & Papenmeier, F. (2021). From metacognitive beliefs to strategy selection: does fake performance feedback influence cognitive offloading? *Psychological Research*, *85*, 2654–2666.

Hertz, N., & Wiese, E. (2019). Good advice is beyond all price, but what if it comes from a machine? *Journal of Experimental Psychology: Applied*, *25*(3), 386–395.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive psychology*, *43*(3), 171–216.

Jarvis, A. L., Hoggan, B. L., & Temby, P. (2022). Neurotracker multiple object tracking ability predicts novice performance on a simulated air traffic control task. *The International Journal of Aerospace Psychology*, *32*(4), 165–182.

Jeffreys, H. (1998). *The theory of probability*. OuP Oxford.

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse towards algorithms? a comprehensive literature review on algorithm aversion.

Levene, H. (1960). Robust tests for equality of variances. *Contributions to probability and statistics*, 278–292.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.

Pfister, R., Weller, L., & Kunde, W. (2020). When actions go awry: Monitoring partner errors and machine malfunctions. *Journal of Experimental Psychology: General*, *149*(9), 1778.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the finst spatial-index model. *Cognition*, *32*(1), 65–97.

Pylyshyn, Z., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in cognitive sciences*, *20*(9), 676–688.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591–611.

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in cognitive sciences*, *18*(5), 259–267.

Wahn, B., Ferris, D. P., Hairston, W. D., & König, P. (2016). Pupil sizes scale with attentional load and task experience in a multiple object tracking task. *PLoS ONE*, *11*(12), e0168087.

Wahn, B., & Kingstone, A. (2020). Humans share task load with a computer partner if (they believe that) it acts human-like. *Acta Psychologica*, *212*, 103205.

Wahn, B., & Schmitz, L. (2024). A bonus task boosts people's willingness to offload cognition to an algorithm. *Cognitive Research: Principles and Implications*, *9*(1), 24.

Wahn, B., Schmitz, L., Gerster, F. N., & Weiss, M. (2023). Offloading under cognitive load: Humans are willing to offload parts of an attentionally demanding task to an algorithm. *Plos one*, *18*(5), e0286102.

Weis, P. P., & Kunde, W. (2024a). Perseveration on cognitive strategies. *Memory & Cognition*, *52*(3), 459–475.

Weis, P. P., & Kunde, W. (2024b). When feedback backfires: Knowledge of results can impair cognitive strategy choice. *Quarterly Journal of Experimental Psychology*, 17470218241282659.

Weis, P. P., & Wiese, E. (2019). Problem solvers adjust cognitive offloading based on performance goals. *Cognitive Science*, *43*(12), e12802.

Weis, P. P., & Wiese, E. (2022). Know your cognitive environment! mental models as crucial determinant of offloading preferences. *Human factors*, *64*(3), 499–513.

Wiese, E., Weis, P. P., Bigman, Y., Kapsaskis, K., & Gray, K. (2022). It's a match: Task assignment in human–robot collaboration depends on mind perception. *International Journal of Social Robotics*, *14*(1), 141–148.