UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXVII CICLO DEL DOTTORATO DI RICERCA IN

NEUROSCIENZE E SCIENZE COGNITIVE

STOCKHOLM UNIVERSITY

COGNITIVE OFFLOADING IN ROUTE PLANNING

Settore scientifico-disciplinare: **M-PSI/01**

DOTTORANDO / A
IRENE FLOREAN

COORDINATORE
PROF. ANDREA CARNAGHI

SUPERVISORE DI TESI
PROF.SSA DONATELLA FERRANTE

CO-SUPERVISORE DI TESI
PROF. FABIO DEL MISSIER

ANNO ACCADEMICO 2023/2024

UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXVII CICLO DEL DOTTORATO DI RICERCA IN

NEUROSCIENZE E SCIENZE COGNITIVE

STOCKHOLM UNIVERSITY

COGNITIVE OFFLOADING IN ROUTE PLANNING

Settore scientifico-disciplinare: **M-PSI/01**

DOTTORANDO / A
IRENE FLOREAN

COORDINATORE
PROF. ANDREA CARNAGHI

SUPERVISORE DI TESI
PROF.SSA DONATELLA FERRANTE

CO-SUPERVISORE DI TESI
PROF. FABIO DEL MISSIER

ANNO ACCADEMICO 2023/2024

Summary

ABSTRACT

Cognitive offloading refers to "the use of physical actions to alter the information processing requirements of a task so as to reduce cognitive demand" (Risko & Gilbert, 2016). The present thesis aims to broaden research on cognitive offloading by investigating this phenomenon in the previously unexplored cognitive domain of route planning. The studies presented in this thesis also seek to deepen our understanding of the consequences of cognitive offloading on task performance.

Cognitive offloading has been extensively investigated in relation to perception, short-term memory, and delayed intentions. In these domains, cognitive offloading has been consistently found to improve task performance in terms of accuracy and/or speed (e.g., Gilbert, 2015a). However, even in the absence of performance advantages, individuals tend to rely on cognitive offloading after its initial use (Gilbert et al., 2020; Scarampi & Gilbert, 2020). Among the predictors of cognitive offloading explored so far, higher cognitive load associated with the task (Risko et al., 2014), lower cognitive abilities (Gilbert et al., 2020), reduced metacognitive confidence in unaided performance (Gilbert et al., 2020), and the avoidance of cognitive effort (Sachdeva & Gilbert, 2020), appear to play a significant role in determining the use of cognitive offloading over internal resources.

Despite providing important insights, research in this area has generally employed relatively simple tasks predominantly focused on memory-related and perceptual domains, leaving cognitive offloading in more complex tasks and domains largely unexplored. Study 1 aimed to address this gap by investigating cognitive offloading in the context of a complex spatial planning task. Specifically, the study examined which offloading strategies were spontaneously devised by individuals while planning routes on maps, as well as the generalizability of findings from previous research on cognitive offloading to this task. In particular, Study 1 explored whether using offloading strategies improved route planning performance, whether greater task difficulty led to a greater reliance on these strategies, and whether they continued to be used even in the absence of performance advantages. The study also investigated the relationship between offloading strategy use and individual differences in cognitive abilities relevant to the task, as well as participants' subjective perceptions of task difficulty and effort.

Building on findings from Study 1, Studies 2 and 3 aimed to extend the understanding of the consequences of cognitive offloading on subsequent task performance. While previous research on cognitive offloading has primarily focused on the impact of external aids (e.g., the Internet, digital storage) on episodic memory (Sparrow et al., 2011; Henkel, 2014), less is known about how engaging in cognitive offloading influences future performance on the same task. Studies 2 and 3 sought to address this gap by investigating what occurs when the opportunity to offload cognition is removed or introduced in a subsequent stage of a route planning task. Both studies hypothesized asymmetric transfer effects, predicting a greater decline in performance following the removal of offloading opportunities than the improvement observed when offloading was introduced, due to asymmetric effects of shifts in planning strategies. Similar to Study 1, Study 2 focused on offloading strategies spontaneously devised by participants, while Study 3 employed external visual aids to support the cognitive offloading strategies identified in the first two studies.

Finally, Study 4 investigated whether the findings from the previous studies could be replicated when using more ecological materials. Specifically, the study assessed whether the external aids employed in Study 3 on fictitious maps for supporting the identified offloading strategies remained effective in supporting route planning when embedded in a real-world map.

By addressing these questions, the present thesis aims to contribute to the literature on cognitive offloading by investigating its role in a more complex cognitive task and previously overlooked domain. Furthermore, this work seeks to deepen our understanding of the consequences of cognitive offloading on task performance when the opportunity to offload cognition is no longer available. Finally, the thesis examines the generalizability of the findings when real-world materials are used.

INTRODUCTION

Imagine that you are visiting a new city for a short time and have a long list of interesting places you would like to see. You need to plan carefully the order in which to visit each place to reduce unnecessary transfers, and you also need to consider some constraints, such as the opening times of places. In other words, you have to find the shortest path to take from a starting point (e.g., your hotel) to an arrival point (e.g., the restaurant chosen for dinner), while stopping by a series of other intermediate points and considering some ordering constraints. You will probably examine a map, look carefully at the position of each place you will visit and find the shortest path connecting them, while respecting constraints. If you wish to see only a few places, the task will be relatively easy, but as the number of places increases, you will probably find yourself marking out the start and end points of your journey on the map, highlighting places to be visited, marking out possible constraints, and tracking your planned path. Using a pen or any other device to facilitate this route planning task is a good example of external cognitive offloading, and the task described in this scenario is very similar to the route planning task we investigated in the studies presented in the following sections.

Cognitive offloading has been defined as "the use of physical actions to alter the information processing requirements of a task so as to reduce cognitive demand" (Risko & Gilbert, 2016, p. 676). While cognitive offloading has been mainly studied in perception, short-term memory, intention offloading, and transactive memory systems (Risko & Gilbert, 2016, see section 1.1 for a review), other important domains, such as route planning on maps unaided by navigation technology, have been neglected. Extending offloading research to uninvestigated cognitive domains contributes to domain-general theories by appraising the robustness of their supportive empirical generalizations and by detecting uncovered aspects that require theory extension or further elaboration. Moreover, the study of uninvestigated domains allows us to discover which offloading strategies are used in these specific domains. This is conducive to the improvement and qualification of cognitive theories within the specific domains and is useful for informing the design of technologies and protocols aimed at supporting domain-specific performance in healthy and weak populations, such as better supportive technology or strategic aids for neuropsychological patients. Before presenting the studies included in this thesis, I will provide a focused introduction to cognitive offloading and to route planning research. The review is selective, including only studies relevant for the goals of the reported studies or for the investigated domain.

1.1 Research on Cognitive Offloading

Various types of cognitive offloading have been observed (Risko & Gilbert, 2016). Offloading onto the body refers to actions that support cognitive tasks by producing physical changes in the body, such as rotating the head to read a tilted text (e.g., Risko et al., 2014) or using the hands to support arithmetic (e.g., Carlson et al., 2007). Offloading into the world, or external offloading, encompasses all the situations in which we modify some aspects of the world to support the accomplishment of cognitive tasks, such as setting external reminders for later retrieval of intentions (Gilbert et al., 2023) and storing information on external memories like notes (e.g., Eskritt & Ma, 2014) and devices (Sparrow et al., 2011). Scholars include in the offloading sphere also the use of rather sophisticated technologies, like GPS, cameras, and search engines (see e.g., Risko & Gilbert, 2016).

Although the specificity of the offloading construct is clearly recognizable, this research stream is conceptually related to other cognitive strands that have examined how cognition stems from a dynamic transaction between the person and the environment in the accomplishment of a task, investigating external representations (e.g., Zhang, 1997), external memory aiding (e.g., Harris, 1980), epistemic actions (e.g., Kirsh & Maglio, 1994), and more generally transactions between artifacts and cognition, including tools and interfaces (Gray et al., 2006; Guthrie et al., 2015; Hutchins, 1995; Lave, 1988; Norman, 1991). These research strands have highlighted the utility of some forms of external support or action to alleviate the cognitive burden in the fulfillment of complex tasks, and/or they have showed how an artifact or an action can change the cognitive operations or strategies involved in the accomplishment of a task.

Several studies investigated intention offloading, defined as "the process of creating a cue in the external environment to trigger a delayed intention. As a result of intention offloading, intentions are stored in our extended physical and social environments, not just in our brains." (Gilbert et al., 2023; p. 61). In their recent summary of existing research on intention offloading, Gilbert et al. highlighted a series of empirical regularities that are relevant for our investigation. Indeed, the authors underlined that intention offloading is used more often when the task is more cognitively demanding, and it improves performance (see also Gilbert, 2015a). However, individuals tend to offload more than necessary, even if this offers no objective benefits in the form of rewards (Fröscher et al., 2022), and to offload more cognition if this was previously allowed vs. not allowed (perseveration effect) (see also Scarampi & Gilbert, 2020). The authors also identified participants' cognitive abilities, metacognitive evaluations, and effort avoidance as potential predictors of the tendency to offload cognition. In particular, participants with lower cognitive ability and lower confidence in their task performance seem to be more willing to offload cognition (Gilbert, 2015b; Gilbert et al., 2020). Moreover, higher extrinsic motivation to sustain effort seems to reduce offloading (Sachdeva & Gilbert, 2020). Indeed, providing financial incentives to promote internal strategies that do not rely on offloading, presumably more effortful, appears to reduce unnecessary offloading behavior (Gilbert et al., 2020).

Research also highlighted potential negative consequences of offloading, such as poorer memory of externally stored contents (Henkel, 2014; Sparrow et al., 2011), overconfidence in one own's knowledge when using search engines (Fisher et al., 2015; Ward, 2021), and long-term weakening of unpracticed abilities (e.g., Carr, 2008; Risko & Gilbert, 2016). Although negative consequences are important and worrisome, considering the widespread diffusion of technologies supporting various forms of offloading, concerns about long-term negative outcomes are currently mainly speculative, and more investigations are needed.

Some studies investigated cognitive offloading also in visuospatial tasks. Moritz et al. (2018) observed that the extraction of information from complex graphical representations was improved when participants had the opportunity to actively change the spatial arrangement of elements of the graphical representations. Evidence also shows that children adaptively resort to external rotation and derive benefits from it when encountering problems that would otherwise necessitate mental rotation for resolution (Armitage et al., 2020). Moreover, in tasks requiring same/different judgments on pairs of differently rotated figures, the tendency to offload cognition through external rotation was found to be sensitive to manipulations of both the actual and the believed reliability of the offloading strategy (Weis & Wiese, 2018), as well as to performance goals (Weis & Wiese, 2019). Overall, these findings confirmed the effectiveness of offloading and its partly adaptive use in relation with task demands.

The relevance of metacognitive evaluations for the decision to offload cognition was highlighted also by Dunn and Risko (2016). Indeed, they found that participants' decisions to tilt their heads to read rotated words were aligned with their judgments on the perceived costs and benefits of external rotation and mental rotation, rather than with the actual performance costs and benefits associated with each type of rotation.

Another aspect that deserves to be mentioned in relation to the spatial domain, and considering the domain of our studies, is the use of GPS technology as an extreme form of offloading. GPS and navigation technology can provide routes to reach target locations and effectively guide individuals in following a path with little human intervention. However, some evidence suggests that offloading wayfinding onto GPS technology may impair spatial memory (e.g., Fenech et al., 2010; Gardony et al., 2013; for a review, see Brügger et al., 2019) because focusing on the device may imply paying less attention to the surroundings (Gardony et al., 2015), potentially leading to poorer spatial memory and subsequent wayfinding if the technology is unavailable. However, further research is needed to consolidate the existing evidence. According to some scholars, intensive use of GPS technology could even impair spatial memory in the long term (Dahmani & Bohbot, 2020), although the available evidence is too limited to draw any firm conclusion. Interestingly, navigation technologies that promote learning, encouraging users to integrate spatial knowledge into a cognitive map, seem able to improve navigation while preserving spatial knowledge and memory (Liu et al., 2022). Better outcomes can be obtained also with navigation technologies in which the level of automation is reduced to promote a more attentive and self-directed encoding of environmental elements (Brügger et al., 2019). These studies indicate that it could be wise to consider less automated forms of assisted planning and navigation to support learning of cognitive maps (Epstein et al., 2017; Thorndyke & Hayes-Roth, 1982), which may be essential if navigation technology is temporarily unavailable, and, more in general, to preserve human spatial memory and skills. However, to inform the design of these technologies from a user-centered prospective, we need to know more about the ways in which individuals interact with external maps while planning routes, and how they offload route planning on these maps.

Summarizing, the study of cognitive offloading has captured significant attention in the last two decades, but research has mainly examined offloading in perception, short-term memory, memory for delayed intentions, some aspects of spatial reasoning, and transactive memory systems (Risko & Gilbert, 2016), while offloading in other important domains, such as route planning on maps unaided by navigation technology, has been overlooked. The results of these investigations have shown communalities and differences, indicating regularities generally holding across domains in some aspects of cognitive offloading, while other aspects seem to be domain specific. Among the more general aspects, it is usually found that cognitive offloading improves task performance, that it is employed more frequently when the task is more cognitively demanding, and that it tends to be used more than necessary for obtaining accurate results. Furthermore, individuals' metacognitive evaluations of the benefits and costs associated with using offloading strategies (vs. not using them) (Risko & Gilbert, 2016), as well as individuals' confidence in their own unaided abilities (Gilbert, 2015b), seem to play a significant role in their decision of whether to offload cognition or not.

The findings highlighting domain specificity refer to the peculiar offloading strategies used by participants in different domains (e.g., tilting the head, setting external retrieval cues in advance, rotating external stimuli) and to the inconsistent findings in participants' tendency to offload cognition across different cognitive tasks (Meyerhoff et al., 2021). This signals that domain-specific and/or task-specific factors are probably involved rather than a single common factor underlying individual differences in offloading behavior across different domains and tasks.

1.2 Route Planning

Wiener, Büchner, and Hölscher (2009; see also Dalton et al., 2019) presented a taxonomy of wayfinding tasks aimed at promoting a better understanding of the cognitive processes involved in their accomplishment. The taxonomy classifies tasks according to various criteria, including the presence of supporting aids, the goal of reaching specific locations, existing knowledge about the surrounding environment and appropriate routes. The *aided wayfinding* category includes tasks in which wayfinding is supported by some kind of external aid, like a map. The route planning task we used in the studies presented here (see, e.g., section 2.2.4.1) can be considered as an instance of this category.

Wiener et al. (2009) assume that external aids support critical cognitive processes for wayfinding, such as planning, decision-making, learning, and memory. However, their taxonomy is not explicit on how the availability of a map actually supports these processes. Moreover, the possibility to act on the map, offloading cognition to gain further support for planning and memory, is not explicitly considered. We investigated this possibility in our studies, in which individuals had to plan a complex route in advance (Montello & Sas, 2006), using a map and exploring various points of interest in an unfamiliar environment and respecting planning constraints. Given the nature of our studies, our review of route planning will focus on studies in which an individual was asked to plan a path on a map or other 2D space to reach one or more destinations from a starting point, in some cases also satisfying given constraints. However, in the final part of this section, we will also consider relevant insights from studies on wayfinding in a broader context.

An influential and seminal contribution to route planning on maps was provided by the work of Hayes-Roth and Hayes-Roth (1979), which found that participants relied on both strategic and opportunistic forms of planning in their multiple-errands tasks. Indeed, participants combined more general top-down strategies with bottom-up hints from the observation of local environmental features. Another important insight provided by this work and the work on planning and problem solving in general (cf. Morris & Ward, 2004) is that individuals plan incrementally and rarely produce complete plans in advance, at least in moderately complex to difficult problems (see also Altmann & Trafton, 2002; Newell & Simon, 1972; Phillips et al., 2001). Empirical studies and simulations provided evidence for this incremental approach also in path planning (e.g., Fum & Del Missier, 2000a; Reitter & Lebiere, 2010). These studies also reported participants' heuristic tendency to maximize the length of straight paths leading from the current position to the next subgoal (see also Bailenson et al., 1988). Other evidence suggests that when individuals are required to find the shortest path among predetermined locations, they tend to prefer routes that minimize boundary crossings between perceived regions of the environment, rather than choosing alternative routes of the same length. This tendency has been linked to the hierarchical planning heuristic (Wiener & Mallot, 2003; see also Wiener, Ehbauer, & Mallot., 2009). However, scholars also pointed out that multiple-errands or path planning tasks can be solved with different strategies in different conditions (e.g., Hayes-Roth, 1979; Fum & Del Missier, 2000b), varying significantly in the demands they pose on cognitive planning vs. visual routines (Hayes-Roth & Thorndyke, 1979).

Another significant thread of studies on spatial planning in 2D spaces deals with human performance in the Travelling Salesman Problem (TSP) and its variants.[1] These studies showed that the TSP, despite its computational complexity, can be solved very effectively by humans using visual strategies (MacGregor & Chu, 2011; MacGregor & Omerod, 1996). The quality of the solutions is generally very close to the optimal one (i.e., the shortest path), although it tends to decrease when the complexity of the problem increases (e.g., more nodes to be connected or presence of complex obstacles). One of the prevailing explanations for this surprising ability is that individuals rely on very

[1] A standard formulation of the TSP requires participants to find an itinerary connecting a given set of *n* cities with specified distance (or cost) between them so that: (1) each city is visited only once, (2) the total distance (or cost) is minimized. The TSP is computationally very complex, belonging to the class of NP-complete problems.

effective visuo-spatial procedures to deal with ordinary perceptual stimuli, and these procedures can be seamlessly applied to TSP (e.g., MacGregor & Ormerod, 1996), although other scholars postulate that higher-level cognitive processing also contributes to participants' performance (Vickers et al., 2001), possibly building over visuo-spatial procedures (van Rooij et al., 2006). Indeed, individual differences in both visuo-spatial abilities (e.g., Burns et al., 2006) and high-level cognitive skills (e.g., Vickers et al., 2001, 2004) have been related with TSP performance, with high-level cognitive skills being more relevant in more complex versions of the TSP (Chronicle et al., 2008).

Some findings from wayfinding research suggest that familiarity with the environment and the nature of the planning task are important factors to consider in route planning (Muffato & Meneghetti, 2020). Indeed, familiarity with landmarks and perceived sense of direction foster shortest path finding in familiar environments (Muffato et al., 2021; Muffato & Meneghetti, 2020). If the shortest route is determined in a real and familiar environment, but without the aid of a map, different routes will indeed result depending on whether they are planned during actual navigation (situated planning) or for future navigation. Shorter routes are designed in situated planning, where spatial information can be cued or provided by the environment, as opposed to prospective planning, where it must be retrieved from a mental representation (Muffato et al., 2021; Hölscher et al., 2011). However, this also suggests that the external availability of spatial information can support route planning by offloading some components of the task (e.g., visually recognizing locations and routes instead of retrieving this information). The availability of a map can play a similar role by making spatial information externally available for route planning.

Finally, individual differences in spatial working memory were found to correlate positively with the accuracy of both situated and prospective planning of the shortest path in familiar environments (Muffato et al., 2021). Wiener, Ehbauer, and Mallot (2009) also referred to spatial working memory to explain their results in a navigational adaptation of the TSP. In their task, participants had to navigate the shortest path between a subset of several boxes placed on the floor in a room. In one group the target boxes to be included in the solution were visually marked with a black marker, while in another group they were indicated in a list. Wiener et al. observed that the navigated routes were shorter in the first group when the number of target boxes was larger. Although working memory was not measured, the authors argued that this difference between the groups was due to the different load on spatial working memory. Indeed, in this study, the visually marked target locations were directly accessible to perception, which facilitated the application of global planning strategies. In contrast, when target locations were not immediately recognizable, they had to be localized and their positions maintained in spatial working memory for planning. In the latter case, when the number of targets exceeded working memory capacity, participants needed to resort to simplified but less effective planning heuristic.

STUDY 1.

SPONTANEOUS COGNITIVE OFFLOADING IN ROUTE PLANNING[2]

2.1. Introduction

2.1.1 Main Aims and Hypotheses

The main aim of our Study 1 was to investigate the spontaneous use of external offloading strategies in route planning on maps. Indeed, the initial research on offloading in a new domain and on a new task requires an understanding of which strategies are spontaneously used and when. Our second aim, in line with previous research (e.g., Gilbert, 2015a), was to understand whether offloading provides performance benefits, especially when the task is more difficult. Finally, we were interested in understanding the relationships between route planning on maps with/without offloading and individual differences in visual search and spatial working memory.

2.1.1.1 Which Offloading Strategies Are Spontaneously Used?

Our route planning task required participants to find the shortest path from a starting place to an arrival place, visiting intermediate locations, and following constraints on the order in which some places had to be visited (see section 2.2.4.1). The order constraints were included to make the task more realistic, following previous route planning studies (e.g., Hayes-Roth & Hayes-Roth, 1979; Wilson et al., 1996). After a cognitive task analysis and preliminary testing, we specified *a priori* four main types of offloading strategies we expected to observe in our task. These strategies were: (1) marking out start and arrival locations; (2) marking out intermediate locations to be visited; (3) marking out ordering constraints; (4) keeping track externally of the progress in the accomplishment of the task (for instance, by drawing the path on the route while planning).

[2] Study 1 has been published. Florean, I., Stragà, M., Mäntylä, T., & Del Missier, F. (2024). Mark it out! Spontaneous cognitive offloading in route planning. *Thinking & Reasoning*, 1-38. https://doi.org/10.1080/13546783.2024.2413151

Indeed, to plan a path starting from one location and ending in another location on a map, with intermediate visits to other locations along the way, participants first need to identify the positions of the target locations on the map, and then remember them. Marking out their positions on the map is a way to help visual search and memory by having these locations popping out against distractor places (Wolfe & Horowitz, 2017, see also Wiener, Ehbauer, & Mallot, 2009). Moreover, if participants must satisfy ordering constraints while planning their paths, externalizing these constraints on the map could be a way to strengthen participants' memory of the intentions to follow these constraints (Gilbert et al., 2023). Finally, participants may need to keep track of the progression of their work and remember the final route for reporting it and, within this process, tracking an external path or sequence is useful both for placekeeping (i.e., remembering in which stage of the planning process they are, see Hambrick, & Altmann, 2015) and for remembering the route when the plan is completed and the solution needs to be reported (Schönpflug, 1986).

Following our cognitive task analysis and preliminary testing, we developed a coding scheme for offloading strategies before data collection and employed it in our study to identify and score the strategies used by our participants (see section 2.2.4.1 and Appendix A, section 1 and Table A.A.1).

2.1.1.2 When Are Offloading Strategies Used?

In this study, we manipulated the possibility to offload cognition by allowing (vs. not allowing) participants to use the pen on the map sheet during route planning (offloading vs. no offloading), the difficulty of the task (low vs. high difficulty maps), and the order in which the maps, varying in difficulty, were presented (low vs. high difficult maps first) (see also section 2.2.2).

Our first hypothesis was that participants in the offloading condition would employ more types of offloading strategies in the high (vs. low) difficulty maps (H1), given that offloading cognition is specifically useful in more complex problems associated with a higher cognitive and memory load (Gilbert, 2015a; Gilbert et al., 2023).

We were also interested in appraising the replicability of the offloading perseveration effects observed in other cognitive domains (e.g., Scarampi & Gilbert, 2020). Accordingly, our second hypothesis was that participants in the offloading condition would use more types of offloading strategies in the low difficulty maps when these maps were presented after the high difficulty maps, promoting offloading use, vs. when they were presented first (H2).

2.1.1.3 Does the Possibility to Offload Cognition Affect Performance and Perceived Effort?

Our third hypothesis was that participants in the offloading (vs. no offloading) condition would show a better planning performance in the high difficulty maps but not in the low difficulty maps (H3). Indeed, if the task is easier, participants' cognitive resources would suffice to plan routes effectively without aiding themselves by offloading cognition, but this would not be the case if the task is more difficult.

An additional hypothesis related to the use of offloading strategies was that participants in the offloading condition would perceive the task as less effortful than participants in the no offloading condition (H4) given that offloading allows to externalize cognition and reduce cognitive and memory load, with the reduction of cognitive effort assumed as one of the drivers of the tendency to offload cognition (Gilbert et al., 2023).

2.1.1.4 What Are The Relationships Between Individual Differences in Cognitive Abilities and Route Planning on Maps With/Without Offloading?

Finally, in the light of previous studies (see Meyerhoff et al., 2021), we were interested in understanding the relationships between individual differences in cognitive abilities and route planning on maps with/without offloading. We focused on visual search and spatial working memory, as these skills appear to be especially relevant in our spatial domain (e.g., Boot et al., 2009; Shah & Miyake, 1996) and in particular in wayfinding/route planning tasks (e.g., Muffato et al., 2021; Wiener, Ehbauer, & Mallot, 2009), while also controlling for the potential influence of metacognitive expectations on performance with reference to the task at hand, as in previous studies on cognitive offloading (Gilbert, 2015b).

Following previous research (Gilbert, 2015b; Gilbert et al., 2020, Meyerhoff et al., 2021), we expected that participants in the offloading condition would use more types of offloading strategies to support their planning processes when they had lower scores on the visual search and working memory measures (H5). We also expected that participants in the no offloading condition would show better planning performance on the high difficulty maps when they had higher scores on the visual search and visual working memory measures (H6), considering that their planning processes were not externally supported (Meyerhoff et al., 2021) and would require greater involvement of working memory and visual search skills.

2.2. Method

2.2.1 Participants

An a-priori power analysis indicated a minimum sample size of 100 participants to detect medium effect sizes for our experimental design ($\eta_p^2 = 0.06$, $f = 0.25$, $\alpha = 0.05$, $1 - \beta = 0.95$). We recruited 113 undergraduates for the study. Participants who failed to comply with the instructions of the planning task ($n = 2$), those with significant handwriting difficulties due to motor deficits ($n = 1$), and those whose performance deviated by more than three standard deviations from the sample mean on at least three measures of the planning task ($n = 1$) were excluded from the analyses. Therefore, the final sample included 109 participants (83% female, age: $M = 22.31$, $SD = 5.85$). The study was approved by the Ethical Committee of the University of Trieste, and participants provided their informed consent prior to participating.

2.2.2 Experimental Design

The study followed a 2 (condition: offloading *vs.* no offloading) × 2 (map difficulty: low *vs.* high) × 2 (presentation order of maps: high difficulty → low difficulty *vs.* low difficulty → high difficulty) mixed experimental design. The possibility of using spontaneous offloading strategies during the route planning task was manipulated between participants, who were told by the instructions that they could freely use a pen on the map sheets if they wanted (offloading condition) or that they could use the pen only to write down their plan in the appropriate response frame on the map (no offloading condition). The difficulty of the planning task was manipulated within participants by increasing the number of places to be visited (and the distractors) in the maps and the number of ordering constraints to be satisfied (see section 2.2.4.1). The low *vs.* high difficulty maps were presented in two separate blocks, and the block order was counterbalanced to detect potential perseveration effects of offloading strategies from the first block to the second (high difficulty → low difficulty *vs.* low difficulty → high difficulty).

2.2.3 Procedure

After reading and accepting the informed consent, participants were randomly assigned to the offloading (n = 55) or no offloading (n = 54) conditions and they had to complete the route planning task. Participants read the instructions for the first block of maps, completed a practice map, and provided a judgment on their expected planning performance in the upcoming block of maps (whose difficulty level was the same as the just-completed practice map). Once finished with the first block of maps, participants provided a retrospective judgment of their planning performance and indicated how difficult and effortful completing the block was. The same sequence (instructions, practice map, performance prediction, route planning block, retrospective judgment of performance, and evaluation of perceived difficulty and effort) was followed in the second block of maps, whose difficulty level was the complement of the first. Afterwards, participants carried out a computerized version of the visual search task (MFTC; Marra et al., 2013), completed demographic measures, and performed the Symmetry Span Task (Kane et al., 2004). The entire procedure lasted approximately 70-80 minutes. After the completion of the data collection and before the debriefing, we invited participants to fill out a short on-line questionnaire, including one question asking them to indicate what they thought the purpose of the study was. This question was asked to check the potential demand effect of our offloading manipulation (i.e., the reference to the use of the pen in the instructions; see also Coupey, 1994; Mandel & Johnson, 2002). Seventy-eight participants responded, stating that the experiment was aimed at investigating visual-spatial, memory, reasoning, planning, and/or decision-making abilities. None mentioned the use of the pen or its potential impact on task completion. Full disclosure of the study aims and additional information were provided in a debriefing session after the completion of data collection.

2.2.4 Materials

2.2.4.1 Route Planning Task

Our paper-and-pencil route planning task presented a series of maps of abstract locations, requiring participants to find the shortest possible path from a given starting place to a given arrival place, visiting specific intermediate locations, and following prescribed ordering rules[3].

The stimuli were designed to study offloading behavior during route planning while minimizing potential confounds due to familiarity with the places (Muffato & Meneghetti, 2021) or peculiar features of the stimuli (e.g., prominent landmarks of known categories like church, city hall, etc.). To this aim, we used maps of artificially created environments instead of real maps. For the same reasons, and to control for potential semantic/lexical confounds (see, e.g., Wiener & Mallot, 2003), locations on the maps were labelled with single letters instead of more complex and familiar names (e.g., church).

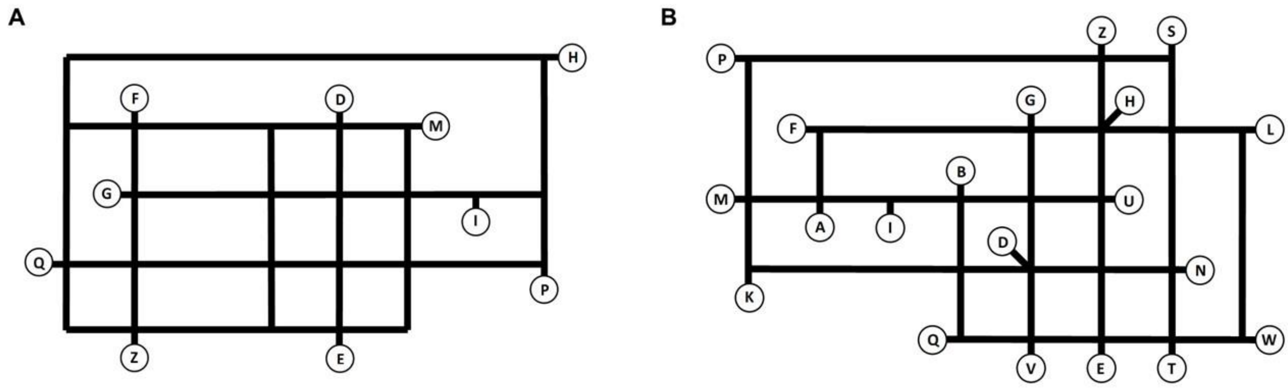
We initially created 24 high difficulty maps. Their structure was generated by randomly selecting 20 out of 40 points from a 8 × 5 rectangular matrix of points, with adjacent points connected by horizontal and vertical segments. The unselected points and their connecting segments were then removed. Twenty letters from the Latin alphabet were randomly chosen and associated with the selected points, following a left-to-right, top-to-bottom sequence. For each map created in this way, a

[3] The task we used is inspired by the Zoo Map of the Behavioral Assessment of Dysexecutive Syndrome battery (Wilson et al., 1996). The Zoo Map seems to be sensitive to aging (Allain et al., 2005) and to planning deficits in schizophrenia (Holt et al., 2013), and it showed to be a good measure of planning in clinical samples (Oosterman et al., 2013).

starting point, an endpoint, and eight intermediate points to be visited were randomly selected. Thus, each high difficulty map included 10 locations to be visited and 10 distractor locations. Two ordering constraints for each high difficulty map were specified by randomly selecting two pairs of locations among the ones to be visited in that map, with the exclusion of the starting and endpoints, and randomly deciding the order of visit (e.g., visit E immediately before visiting W *vs*. visit to W immediately before visiting to E). Twenty-four low difficulty maps were generated from the high difficulty maps. From each of the original 24 high difficulty maps, we retained, in its low difficulty counterpart, the starting and arrival points, and the following randomly selected elements: one ordering constraint with the related locations, one additional location to be visited, and half of the distractor locations. The segments connecting the unselected locations were eliminated. Therefore, each low difficulty maps included a starting point, an arrival point, three intermediate locations to be visited, five distractor locations, and one ordering constraint (for examples see Figure 1).

Figure 1

Examples of a Low Difficulty Map (Panel A) and a High Difficulty Map (Panel B)

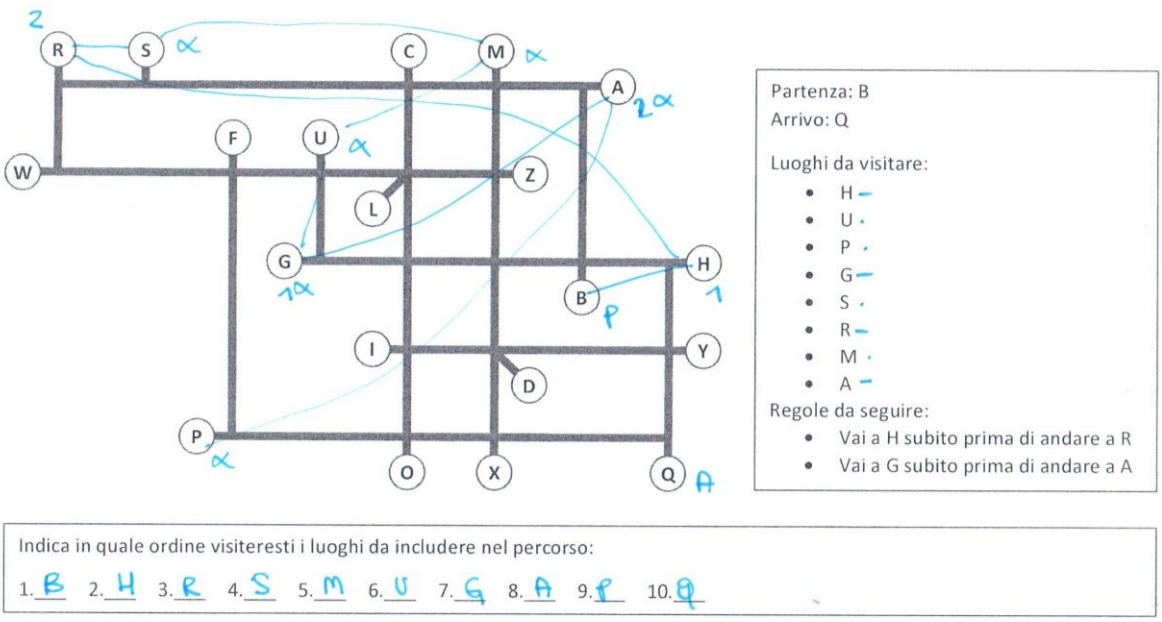


After a pretest, a total of 16 high difficulty maps and 16 low difficulty maps were selected to create two sets of maps of similar difficulty, with each set including eight high difficulty maps and eight low difficulty maps. Each set included a given map either in its high difficulty version or in its low difficulty version. Participants were presented with only one of the two set of maps, in two blocks (one comprising eight high difficulty maps and the other eight low difficulty maps), with set assignment counterbalanced among participants in both experimental conditions (i.e., offloading vs. no offloading). To avoid order effects, maps were presented in random order within each block.

Participants were given a booklet of layered sheets, each containing a map and its specific directions in the rightmost area (i.e., starting and arrival places, locations to be visited, rule/s to follow). Participants were instructed to identify the shortest path as quickly and accurately as possible and, only once they had identified their solution in its entirety, to indicate it within a designated response frame below the map. For this purpose, they had to write down all the letters (locations) of their planned path exactly in the planned order of visit (for an example of the complete map sheet, see Figure 2).

Figure 2.

High Difficulty Map Completed by a Participant in the Offloading Condition



Note. The rightmost instructions frame presents the starting point (B – *Partenza*), the arrival point (Q – *Arrivo*), the places to be visited (H, U, P, G, S, R, M, A – *Luoghi da visitare*), and the two ordering constraints (visit H immediately before visiting R; visit G immediately before visiting A – *Regole da seguire*). The response frame below the map is where the planned route had to be reported. The figure shows a participant's use of four external offloading strategies (marking out starting and arrival locations, marking out intermediate locations to be visited, marking out ordering constraints, and track keeping).

The instructions for all the participants were identical, except for a single difference between the offloading and the no offloading conditions. Participants in the offloading condition were simply informed that if they wanted, they could freely use the pen on the map sheet even before identifying the final route. Conversely, participants in the no offloading condition were informed that they had to use the pen exclusively to write the final route in the response frame. After writing down the entire route, participants had to flip the sheet and place it in a box on their right to proceed with the next map.

To assess accuracy of performance, the shortest path for each map among the ones completed without errors by all participants was identified (i.e., including the starting and arrival places and the locations to be visited, and respecting the ordering rule/s). Performance for each map solved without errors was measured as the proportional deviation above the shortest route, namely the proportion by which the planned route length exceeded the shortest route for that map (see, e.g., MacGregor & Ormerod, 1996), computed as follows: *proportional deviation* = (actual route length - shortest route length) / shortest route length. This measure reflects the quality of the planned route, with higher values indicating a greater deviation from the shortest route and zero indicating no deviation. We used this relative evaluation of performance because, to date and to the best of our knowledge, no automated algorithm is available for computing the shortest route for visiting a series of places when the order of visits of some of them is constrained. Another measure of accuracy was the proportion of maps completed without errors. To measure the completion time of a given map, the time interval between the flip of the previous map sheet and the flip of the given map sheet was recorded. Experimenters also unobtrusively recorded the time in which participants started and ended writing the solution within the response frame for each map. The time measures were computed on the maps completed without errors.

2.2.4.2 Coding and Scoring of Offloading Strategies

To classify the offloading strategies spontaneously devised by participants, each map completed in the offloading condition was examined by one expert rater (one of the authors) reporting the use in the map of each of the four categories of offloading strategies specified in the task analysis: (1) marking out start and arrival locations; (2) marking out intermediate locations to be visited; (3) marking out ordering constraints; (4) track keeping. Therefore, each map completed by a given participant was scored with a 0-1 dichotomous measure for each offloading strategy, assigning 0 when the strategy was not implemented, and 1 when it was. A summative score, representing the total number of strategies applied by each participant for each map, was also computed. The score ranges from zero (no offloading strategy employed) to four (all offloading strategies employed). Subsequently, the number of strategies computed individually for each map was averaged across maps of the same difficulty level to derive a measure of offloading strategy use for each difficulty block. To compute inter-rater reliability, two trained raters examined 20% of the maps completed in the offloading condition and classified the offloading strategies used. The inter-rater agreement was very good (99% of agreement), with a Cohen's K of .98.

2.2.4.3 Metacognitive Performance Judgments, Perceived Difficulty, and Effort

Performance prediction and postdiction. Following a practice map, participants were asked to estimate how much they thought they would deviate from the shortest possible routes in the upcoming block of maps with the same difficulty level, answering on a 10-point scale ranging from *I will deviate a lot* to *I will deviate very little*. Immediately after completing the first block of maps, participants were again asked to indicate how much they thought they had actually deviated from the shortest possible routes using the same response scale (*I deviated a lot – I deviated very little*). This procedure (adapted from Gilbert, 2015b) was repeated in the second block of maps.

Perceived Difficulty and Effort. Following the completion of each map block, participants were also asked to indicate how difficult the series of maps they had just completed were and how much effort was needed to complete them, using a 7-point response scale from *very little* to *very much*.

2.2.4.4 Measures of Cognitive Abilities

Visual Search. To assess visual search ability, we used a computerized version of the Multiple Features Target Cancellation task (MFTC; Marra et al., 2013). Participants were shown a grid of 80 small squares, each containing two differently oriented lines. They were first asked to fixate a probe (a small square with two oriented lines) placed just above the grid and then to identify which of the 80 squares were identical to the probe (13 targets) by clicking on them as quickly and accurately as possible. Timing started when participants, having read the instructions, entered the screen displaying the target and distractor stimuli, and it stopped when participants pressed a bottom-right arrow on the screen to confirm task completion. The accuracy score was calculated according to the formula indicated by Marra et al. (i.e., {(correct answers / 13) + [(67 – false alarms) / 67]} /2).

Complex Working Memory Span. The computerized version of the Symmetry Span Task (Kane et al., 2004) from the Attention & Working Memory Lab (https://englelab.gatech.edu/) was used to measure spatial working memory (WM) capacity. Each trial presented participants with a series of 8 × 8 matrices of black and white squares and required them to judge whether the matrices were symmetrical on their vertical axis (process component of the task). After each matrix, a red square appeared in a 4 × 4 matrix (storage component of the task). At the end of each series (with a length ranging from 2 to 5), participants were presented with a 4 × 4 empty matrix and they had to recall the correct locations of the previously presented red squares in their order of appearance by clicking on the cells of the matrix. Before performing the actual test trials, participants had to complete three practice sessions. In the first session, they practiced the storage component of the task, in the second the process component, and in the third the full task. During the test trials, if participants spent more time judging symmetry than their average time during the practice session with process trials plus 2.5 SD, the program counted that trial as an error and automatically presented the matrix with the red square. The participants' performance was measured using the partial score rather than the absolute score. The absolute score considers only the trials where the participant recalled all elements in the correct order without errors, whereas the partial score considers all correctly recalled elements within individual trials (Oswald et al., 2015). Previous research has shown that the partial score represents a more reliable and sensitive measure of individual differences in working memory capacity compared to the absolute score (Redick et al., 2012).

2.3 Results

2.3.1 Preliminary Analyses

For each performance variable of the route planning task, we trimmed data within the offloading/no offloading and high/low map difficulty conditions: values that were 3 *SD*s above or below the mean were replaced by the mean ± 3 *SD*s (for a similar procedure, see Del Missier et al., 2021; Miyake et al., 2000). One participant was excluded from the analyses because they deviated more than ± 3 *SD*s from the mean on all the performance measures of the route planning task. For the individual-differences tasks the replacement was done employing the same criterion on the entire sample.[4] Descriptive statistics for all the measures of the route planning task are shown in Table 1.[5]

Table 1

Descriptive Statistics of Route Planning Performance Measures

	Proportion of Maps Without Errors		Map Completion Time		Proportional Deviation Above the Shortest Rath	
Map Difficulty	High Difficulty Maps M (SD)	Low Difficulty Maps M (SD)	High Difficulty Maps M (SD)	Low Difficulty Maps M (SD)	High Difficulty Maps M (SD)	Low Difficulty Maps M (SD)
Condition						
No Offloading	0.85 (0.21)	0.96 (0.07)	186.25 (56.07)	47.28 (13.53)	0.17 (0.13)	0.02 (0.04)
Offloading	0.91 (0.15)	0.95 (0.10)	175.02 (57.99)	52.23 (23.76)	0.10 (0.06)	0.02 (0.04)

Note. Completion times are indicated in seconds.

[4] Data trimming led to the following percentages of value substitution: proportion of high difficulty maps completed without errors: 3.64% for offloading condition and 3.70% for no offloading condition; proportion of low difficulty maps completed without errors: 5.45% for offloading condition and 0.00% for no offloading condition; proportional deviation above the shortest path in high difficulty maps, 3.70% for offloading condition and 3.85% for no offloading condition; proportional deviation above the shortest path in low difficulty maps, 3.64% for offloading condition and 3.70% for no offloading condition; high difficulty map completion time, 1.85% for offloading condition and 1.92% for no offloading condition; low difficulty map completion time, 1.82% for offloading condition and 0.00% for no offloading condition; high difficulty maps solution-writing time, 3.70% for offloading condition and 1.92% for no offloading condition; low difficulty maps solution-writing time, 3.64% for offloading condition, and 1.85% for no offloading condition; 0.92% for visual search (MFTC accuracy scores); 0.00% for working memory (Symmetry Span partial scores). Descriptive statistics for MFTC were as follows: $M = 0.89$, $SD = 0.09$. Descriptive statistics for Symmetry Span were as follows: $M = 27.16$, $SD = 7.57$.

[5] When skewness values of the variables were greater than ± 2 (within the offloading and map difficulty conditions for the route planning performance measures, and in the entire sample for other tasks), we employed a log10 transformation to normalize their distribution. However, there were no significant differences in the results with transformed and untransformed data. Therefore, in the subsequent sections we have reported the analyses on non-transformed data to enhance readability.

No significant differences were found between participants in the offloading and no offloading conditions in terms of gender ($\chi^2 = 1.75$, $p = .186$), visual search abilities ($t(107) = 0.80$, $p = .425$), and spatial working memory ($t(107) = 0.86$, $p = .393$). For descriptive statistics, see Appendix A, Table A.A.2 and Table A.A.3. Correlations between performance measures for all administered tasks are also reported in the Appendix A (Table A.A.4 and Table A.A.5).

As a manipulation check, we tested whether the high difficulty maps were perceived by participants as actually more difficult than the low difficulty ones. A 2 (high vs. low difficulty map block) × 2 (offloading vs. no offloading condition) × 2 (high → low difficulty vs. low → high difficulty presentation order of map blocks) mixed ANOVA on participants' difficulty ratings highlighted significant main effects, with no significant interactions (Fs < 3.55, ps > .06). The perceived difficulty of the task was higher in the high difficulty maps than in the low difficulty ones, ($M_{high} = 4.90$, $SD_{high} = 1.19$ vs. $M_{low} = 2.96$, $SD_{low} = 1.08$), $F(1,105) = 315.73$, $p < .001$, $\eta^2_p = .75$. Participants in the offloading condition perceived the task as less difficult than participants in the no offloading condition ($M_{off} = 3.63$, $SD_{off} = 1.10$ vs. $M_{nooff} = 4.25$, $SD_{nooff} = 1.10$), $F(1,105) = 14.76$, $p < .001$, $\eta^2_p = .12$. Finally, participants who completed the high difficulty maps first perceived the task as less difficult compared to participants who completed the low difficulty maps first ($M_{hl} = 3.59$, $SD_{hl} = 1.07$ vs. $M_{lh} = 4.28$, $SD_{lh} = 1.11$), $F(1,105) = 18.26$, $p < .001$, $\eta^2_p = .15$.

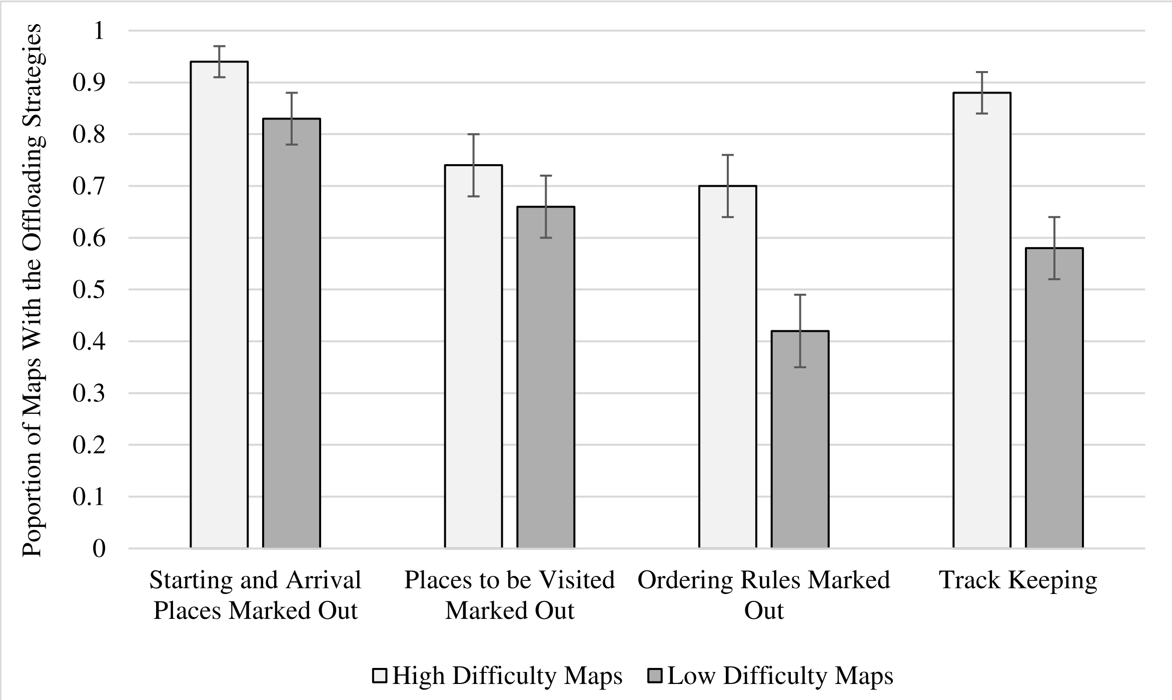
2.3.2 Which Offloading Strategies Are Spontaneously Used and When?

In the offloading condition, we observed a frequent use of all the four strategies we identified, ranging from rule(s) marked out in the 40% of the low difficulty maps to starting and arrival points marked out in the 94% of the high difficulty maps (see Figure 3). Different implementations of the four offloading strategies were observed, but no alternative offloading strategies were identified (see section 2.2.4.2).

To test the hypothesis (H1) that offloading strategies are used more frequently in high difficulty (*vs*. low difficulty) maps, differences in the use of each of the four offloading strategies between the low *vs.* high difficulty maps were appraised with paired-sample t-tests within the offloading condition. A significantly higher average proportion of high difficulty (*vs*. low difficulty) maps were observed with starting and arrival points marked out (M_{high} = 0.94, SD_{high} = 0.23 *vs.* M_{low} = 0.83, SD_{low} = 0.37), $t(54) = 2.25$, $p = .028$, d =.30, rule(s) marked out (M_{high} = 0.70, SD_{high} = 0.43 *vs.* M_{low} = 0.42, SD_{low} = 0.48), $t(54) = 4.94$, $p < .001$, d =.67, and track keeping (M_{high} = 0.88, SD_{high} = 0.31 *vs.* M_{low} = 0.58, SD_{low} = 0.46), $t(54) = 5.16$, $p < .001$, d =.70, whereas no significant differences emerged in marking out places to be visited (M_{high} = 0.74, SD_{high} = 0.43 *vs.* M_{low} = 0.66, SD_{low} = 0.46), $t(54) = 1.57$, p = .123, d = .21 (see Figure 3). In addition, the difference in the overall number of offloading strategies used per map between high *vs.* low difficulty maps was significant (M_{high} = 3.26, SD_{high} = 0.94 *vs.* M_{low} = 2.49, SD_{low} = 1.29), $t(54) = 5.42$, $p < .001$, d = .73. Thus, participants who were free to use the pen during route planning employed more frequently three out of the four offloading strategies we identified in the difficult (*vs*. simple) maps, and they also used overall more types of offloading strategies, in line with H1. However, in the simple maps, participants used an average of 2.49 offloading strategies out of four, marking out starting and arrival points in more than 80% of the maps and places to be visited in more than 65% of the maps. Furthermore, on average participants employed at least one offloading strategy in 86.82% of the low difficulty maps.

Figure 3

Proportion of High Difficulty and Low Difficulty Maps Showing Different Types of Offloading Strategies



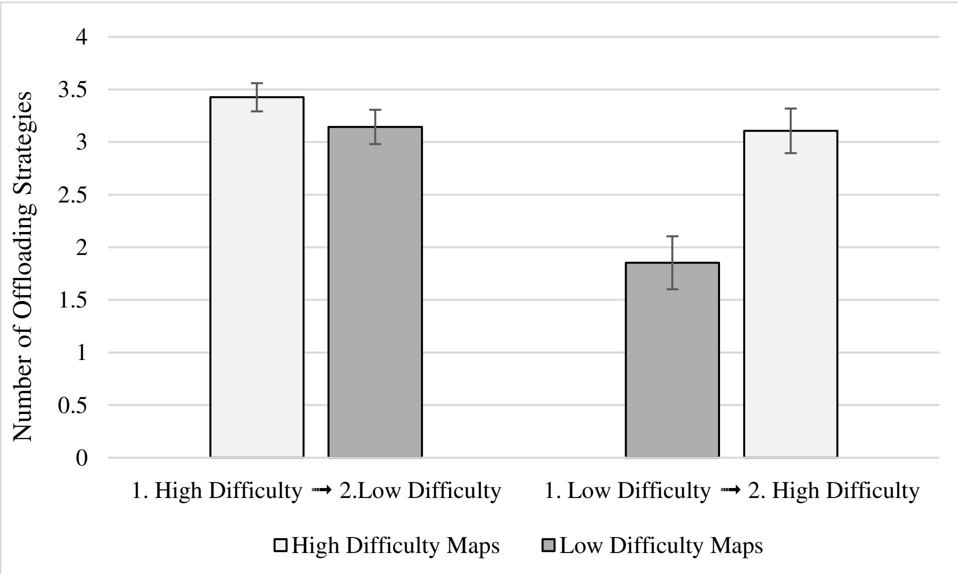
Note. Error bars are standard errors.

To test the hypothesis (H2) that offloading strategies are used more frequently if they had been previously used (perseveration effect), a mixed 2 (low *vs.* high difficulty maps) × 2 (high → low difficulty maps *vs.* low → high difficulty maps) ANOVA was carried out on the average number of offloading strategies employed by participants in the offloading condition. The analysis showed a significant interaction between map difficulty and presentation order, $F(1,53) = 14.31$, $p < .001$, $\eta^2_p = .21$, beyond the main effects of map difficulty, $F(1,53) = 35.78$, $p < .001$, $\eta^2_p = .40$, and presentation order, $F(1,53) = 10.55$, $p = .002$, $\eta^2_p = .17$. Post hoc comparisons showed that participants who completed the high difficulty maps first did not exhibit a significant decrease in the average number of offloading strategies employed when they moved to the low difficulty maps ($M = 3.43$, SD = 0.70 *vs.* $M = 3.14$, $SD = 0.85$), $t(53) = 1.54$, $p = .421$, while participants who completed the low difficulty maps first applied a higher number of offloading strategies when they then moved to the high difficulty maps

(M = 1.85, SD = 1.34 *vs.* M = 3.11, SD = 1.12), $t(53) = 6.97$, $p < .001$ (see also Figure 4). Thus, participants in the former order condition perseverated in using a similar number of offloading strategies despite the decreased difficulty of the maps, in line with H2.

Figure 4

Number of Offloading Strategies as a Function of Map Difficulty and Presentation Order of Map Difficulty



Note. Error bars are standard errors.

2.3.3 Does the Possibility to Offload Cognition Affect Performance and Perceived Effort?

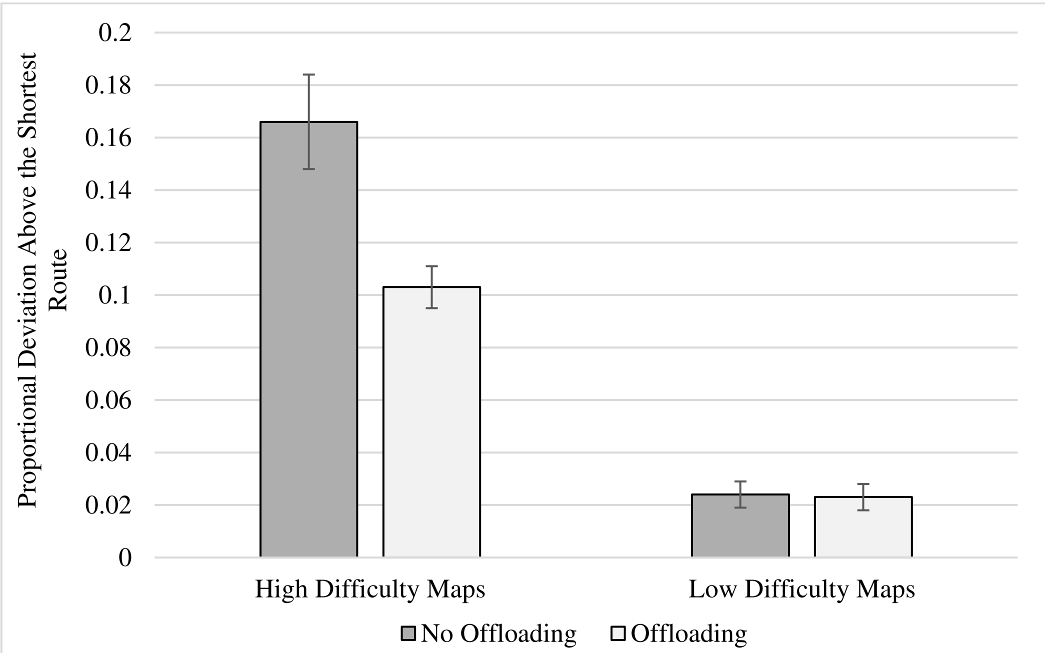
To test the hypothesis that offloading has positive effects on performance in high difficulty maps (H3), a series of 2 (low *vs.* high difficulty maps) × 2 (offloading *vs.* no offloading condition) × 2 (high → low difficulty maps *vs.* low → high difficulty maps) mixed ANOVAs was carried out on each of the performance measures of the route planning task.

Concerning the proportional deviation, the main effects of the offloading condition, $F(1,102) = 8.53$, $p = .004$, $\eta^2_p = .08$, and map difficulty, $F(1,102) = 143.20$, $p < .001$, $\eta^2_p = .58$, were qualified by a significant offloading condition by map difficulty interaction, $F(1,102) = 10.50$, $p = .002$, $\eta^2_p = .09$.

No other effect reached statistical significance (*F*s < 3.12, *p*s > .08). Tukey post hoc tests showed that participants in the offloading condition deviated less from the shortest routes than participants in the no offloading condition in the high difficulty maps (M_{off} = 0.10, SD_{off} = 0.06, 95% CI [0.09, 0.12] *vs.* M_{noff} = 0.17, SD_{noff} = 0.13, 95% CI [0.13, 0.20]), $t(102) = 3.27$, p =.008, but not in the low difficulty maps (M_{off} = 0.02, SD_{off} = 0.04, 95% CI [0.01, 0.03] *vs.* M_{noff} = 0.02, SD_{noff} = 0.04, 95% CI [0.01, 0.03]), $t(102) = 0.29$, $p = .992$; see Figure 5), while both groups of participants in the offloading and no offloading conditions deviated less from the shortest routes in the low difficulty maps than in the high difficulty maps, 95% CI [0.02, 0.03], $t(102) = 6.23$, $p < .001$, and 95% CI [0.11, 0.15], $t(102) = 10.66$, $p < .001$, respectively.

Figure 5

Deviation from the Shortest Route in the Offloading/No Offloading Conditions for the Low Difficulty and High Difficulty Maps



Note. Error bars are standard errors.

A main effect of map difficulty was observed on the proportion of maps completed without errors, $F(1,105) = 16.62$, $p < .001$, $\eta^2_p = .14$, showing a higher proportion of maps completed without errors in the low difficulty maps compared to the high difficulty maps (M_{high} = 0.88, SD_{high} = 0.18, 95% CI [0.85, 0.92] *vs.* M_{low} = 0.96, SD_{low} = 0.09, 95% CI [0.94, 0.97]). All the other effects were not statistically significant (*F*s < 3.28, *p*s > .07).

The mean time needed to complete a map was also affected by map difficulty, $F(1,102) = 752.21$, $p < .001$, $\eta^2_p = .88$, with high difficulty maps requiring more time to be completed than low difficulty maps (M_{high} = 180.53 s, SD_{high} = 57.07 s *vs.* M_{low} = 47.79 s, SD_{low} = 19.81 s). The interaction between map difficulty and offloading condition was statistically significant, $F(102) = 4.23$, $p = .042$, $\eta^2_p = .04$. However, post hoc comparisons did not show statistically significant differences in map completion time between the offloading and no offloading conditions, neither in high difficulty maps, $t(102) = 1.01$, $p = .741$, nor in low difficulty maps, $t(102) = -2.41$, $p = .082$. The interaction between map difficulty and presentation order was also statistically significant, $F(1,102) = 11.04$, $p = .001$, $\eta^2_p = .10$. Post hoc comparisons indicated that participants who completed the high difficulty maps first (*vs*. the low difficulty maps first) spent less time on solving the low difficulty maps, $t(102) = -3.08$, $p = .014$ (M_{hl} = 42.17 s, SD_{hl} = 15.55 s *vs.* M_{lh} = 53.52 s, SD_{lh} = 22.07 s), whereas there were no differences related to the presentation order for the high difficulty maps, $t(102) = 1.91$, $p = .230$ (M_{hl} = 191.15 s, SD_{hl} = 51.96 s *vs.* M_{lh} = 170.31 s, SD_{lh} = 60.30 s). No other effect reached statistical significance (*F*s < 2.23, *p*s > .13).

The opportunity to offload cognition seems to benefit performance in the more difficult maps only, as predicted by H3. In particular, participants who were allowed to offload cognition planned routes deviating less from the shortest ones. However, offloading did not provide statistically significant advantages in terms of error reduction, which was generally low, or time to complete the task.[6]

[6] We further examined the effects of offloading (*vs*. not offloading) on map completion time, carrying out some drill-down analyses to compare offloading and no offloading conditions in each difficulty level within each presentation order. When low difficulty maps were presented first, no differences between participants in the offloading condition *vs*. no offloading condition were found in map completion time, neither on high difficulty maps ($t(50) = 0.15$, $p = .559$, $d = .04$), nor on low difficulty maps ($t(52) = 1.34$, $p = .187$, $d = .36$). However, when high difficulty maps were presented first, participants in the offloading condition needed less time than those in the no offloading condition to complete the high difficulty maps ($t(50) = 1.74$, $p = .044$, $d = .48$). Conversely, participants in the offloading condition needed more time than those in the no

To test the hypothesis that perceived cognitive effort was lower in participants who had the opportunity to offload cognition *vs*. those who did not (H4), we carried out a 2 (low *vs*. high difficulty maps) × 2 (offloading *vs*. no offloading) × 2 (high → low difficulty *vs*. low → high difficulty maps) mixed ANOVA on perceived effort. All the main effects were significant, but none of the interactions were (Fs < 2.65, ps > .10). Participants reported having experienced more effort in the no offloading condition (M_{off} = 3.74, SD_{off} = 1.24 *vs*. M_{nooff} = 4.49, SD_{nooff} = 1.04), $F(1,105) = 19.87$, $p < .001$, η^2p = .16, and in the high difficulty maps (M_{high} = 5.11, SD_{high} = 1.17 *vs*. M_{low} = 3.11, SD_{low} = 1.25), $F(1,105)$ = 273.81, $p < .001$, $\eta^2_p = .72$. Additionally, participants who completed the high difficulty maps first perceived the effort devoted to planning the routes as lower than those who completed the low difficulty maps first (M_{hl} = 3.79, SD_{hl} = 1.12 vs. M_{lh} = 4.44, SD_{lh} = 1.21, $F(1,105) = 14.75$, $p < .001$, $\eta^2_p = .12$). These findings show that offloading cognition on maps is associated with a reduced perceived effort, in accordance with H4.[7]

2.3.4 What Are the Relationships Between Individual Differences in Cognitive Abilities and Route Planning on Maps With/Without Offloading?

We tested our hypotheses concerning individual differences in visual search and spatial working memory with multiple regression and carried out an additional exploratory mediation analysis to provide a more complete account of the results. According to H5, participants in the offloading

---

offloading condition to complete the low difficulty maps ($t(53) = 2.30$, $p = .025$, $d = .62$). These results are in line with the one on the perseveration effect (see section 2.3.2). For a complete report of the analyses, see Appendix A, Table A.A.6 and Figure A.A.1.

[7] Participants who completed the high difficulty maps first perceived the overall task as less difficult (see section 2.3.1) and less effortful (see section 2.3.3) than those who completed the low difficulty maps first. This may be because those in the offloading condition did not significantly decrease their use of offloading strategies when shifting to low difficulty maps (see section 2.3.2). The lack of significant three-way interactions between offloading condition, presentation order, and map difficulty on both perceived difficulty and perceived effort does not support this possibility. However, considering the perseveration effect observed in the use offloading strategies, we decided to further explore this possibility. Thus, we focused on the low difficulty maps completed by participants in the offloading condition, which varied in the number of offloading strategies used with the presentation order, unlike the high difficulty maps. We regressed block presentation order (high difficulty maps first vs. low difficulty maps first) and the number of offloading strategies employed in low difficulty maps on the perceived difficulty of low difficulty maps. None of the examined factors was a significant predictor (βs < |.59|, ps > .061). In the second analysis, we regressed the same predictors on the perceived effort of low difficulty maps. The results showed that block presentation order remained a significant predictor of perceived effort even when controlling for the number of offloading strategies used ($\beta = -.74$, $t(52) = -2.46$, $p = .017$), but the number of offloading strategies was not a significant predictor ($\beta = .07$, $t(52) = 0.44$, $p = .662$).

condition who had lower scores in visual search and working memory measures would use more types of offloading strategies to support their planning processes, in line with previous research on cognitive offloading (Gilbert, 2015b; Gilbert et al., 2020). We entered in the regression model as predictors of the number of offloading strategies used the visual search score (MFTC), the working memory score (Symmetry Span), and two control variables: predictive metacognitive judgments of performance in the upcoming map block and presentation order of the maps (see section 2.2.4). We carried out separate regression for low difficulty maps and high difficulty maps.

In the low difficulty maps, none of the potential predictors was significant, except for the presentation order of the map blocks (in line with H2). In the high difficulty maps, the visual search measure was a significant predictor, with better performance in visual search being associated with a higher number of offloading strategies used (Table 2; regression models on single offloading strategies are reported in Appendix A, Tables A.A.7-A.A.10). These findings did not support H5, showing instead that having better visual search skills is associated with the use of a greater variety of offloading strategies in route planning.

Table 2

Multiple Regression on the Number of Offloading Strategies Used

	High Difficulty Maps					**Low Difficulty Maps**				
Predictors	β	95%CI Lower	95%CI Upper	t	p	β	95%CI Lower	95%CI Upper	t	p
Visual Search (MFTC)	.30	.01	.58	2.08	.043	.18	-.06	.42	1.51	.138
Working Memory (Symmetry Span)	.18	-.08	.45	1.38	.174	.08	-.16	.32	0.66	.514
Metacognitive Prediction	-.02	-.31	.27	-0.16	.870	-.02	-.26	.23	-0.16	.873
Presentation Order of Map Difficulty	-.35	-.88	.18	-1.33	.190	-1.01	-1.49	-.53	-4.21	< .001
Overall Model	Adjusted R^2		F		p	Adjusted R^2		F		p
	.09		2.41		.062	.24		5.31		.001

Note. Lower and upper bounds for 95% confidence intervals of coefficients are indicated with 95%CI Lower and 95%CI Upper, respectively.

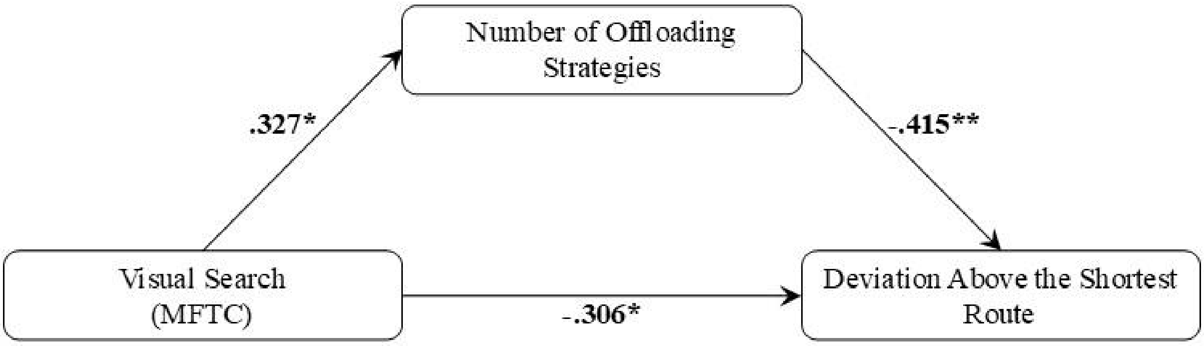
We carried out further exploratory analyses within the offloading condition to understand whether individual differences in cognitive abilities and offloading strategies use were related to planning performance. To this aim, we estimated multiple regression models for the low difficulty maps and for the high difficulty maps with the proportional deviation above the shortest route as the criterion variable. In both regression models, the predictors were visual search score (MFTC), working memory score (Symmetry Span), predictive metacognitive judgment of performance[8], and average number of offloading strategies used. In the low difficulty maps, none of the predictors were significant (βs < .14, *p*s > .37). In the high difficulty maps, individual differences in visual search abilities were negatively related with the deviation from the shortest route, β = -.34, $t(49) = -2.67$, $p = .010$, showing that participants with better visual search skills had a better route planning performance. Moreover, the number of offloading strategies used was negatively related with the deviation from the shortest route, β = -.40, $t(49) = -3.32$, $p = .002$, indicating that applying more offloading strategies supported planning performance. None of the other predictors were significant.

Considering these results, we explored the possibility that the number of offloading strategies used would mediate the relation observed between visual search skills and route planning performance. To this aim, we estimated the model shown in Figure 6. The model shows a partial mediation effect: participants with better visual search skills used more offloading strategies while planning, and this led to a better performance (bootstrapping estimate with 10000 cycles; indirect effect, β = -.136, 95%CI[-.281, -.004]), but the direct effect of visual search skills on planning performance remained significant, β = -.306, $t(51) = -2.56$, $p = .013$. This shows that visual search ability contributes positively to planning performance also beyond its support via the adoption of offloading strategies.

[8]The inclusion of retrospective metacognitive judgments as predictors in the regression models with the proportional deviation as the criterion variable did not change the results reported here.

Figure 6

Model of the Relationships Between Visual Search Skills, Number of Offloading Strategies Used, and Planning Performance in High Difficulty Maps of the Offloading Condition



*$p < .05$, **$p < .01$, ***$p < .001$.

Another set of regression analyses was carried out to test the hypothesis that participants in the no offloading condition with higher scores in visual search and visual working memory measures would show a better planning performance in the high difficulty maps (H6), considering that their planning processes would not be externally supported (Meyerhoff et al., 2021) and high difficulty maps would require a greater involvement of working memory and visual search abilities.

To test this hypothesis, we carried out two multiple regressions, for low difficulty maps and high difficulty maps, using as the criterion variable the deviation from the shortest route, and as predictors the visual search score (MFTC), the working memory score (Symmetry Span), and the predictive metacognitive judgments of performance. In the simple maps, none of the potential predictors were significant (*t*s < 0.72, p*s* > .47). However, in the high difficulty maps, the Symmetry Span score was a significant predictor, $\beta = -.33$, $t(48) = -2.46$, $p = .018$, showing that participants with a higher working memory span had a better planning performance (for the complete report of results see Appendix A, Table A.A.11).

2.4 Discussion

Study 1 aimed to investigate the spontaneous use of offloading strategies in route planning on maps, a domain not previously investigated in relation to cognitive offloading. We wanted to appraise the generalizability of cognitive offloading research, but also to understand the domain-specific aspects of cognitive offloading in route planning on maps.

We first examined the offloading strategies used by participants allowed to use the pen to plan the route on the map. These participants frequently employed all four kinds of offloading strategies defined by an *a priori* task analysis and observed in pilot testing: marking out starting and arrival points, marking out intermediate locations to be visited, marking out ordering constraints, and keeping track of the work done.

Then, we examined when offloading strategies were used, finding that participants in the offloading condition employed more types of offloading strategies in the high (*vs*. low) difficulty maps, in agreement with H1. However, offloading strategies were frequently used also in the low difficulty maps, with no accuracy improvement *vs*. the no offloading condition. Participants in the offloading condition also showed a tendency to perseverate in the use of offloading strategies when they solved the more difficult maps before the easier ones, as predicted by H2.

Next, we appraised the effects of offloading on performance, observing that participants allowed to offload cognition planned shorter routes than participants not allowed to offload, but in the high difficulty maps only (H3). Additionally, participants in the offloading condition perceived a lower effort than those in the no offloading condition (H4).

Finally, we examined the relationships between route planning on maps with/without offloading and individual differences in visual search and spatial working memory. In the offloading condition, we found that a better visual search ability was related with a stronger tendency to offload cognition during route planning in the more difficult maps, contrary to our hypothesis (H5). Further analyses showed that visual search ability was related, directly and through offloading, with better planning performance. In the no offloading condition, participants with better spatial working memory obtained a better planning performance in the more complex maps (H6), but that was not the case for participants with better visual search ability.

2.4.1 Theoretical Implications

Our findings on offloading in the domain of route planning on maps support the generalizability of several results obtained by the research on cognitive offloading in other domains. In particular, in agreement with previous research, we observed: (1) a stronger tendency to offload cognition when the task is more difficult (Gilbert, 2015a; see also Gilbert et al., 2023), (2) the tendency to perseverate in the use of offloading strategies (Scarampi & Gilbert, 2020), and (3) the positive support of offloading to performance (Gilbert et al., 2023). Some of our results also agree with the idea that the reduction of cognitive effort can be one of the factors underlying the tendency to offload cognition (Sachdeva & Gilbert, 2020). Indeed, participants in the offloading condition used offloading strategies also in low difficulty maps with no accuracy benefit (but with a time cost – see note 5, Figure A.A.1, Table A.A.6) compared to participants in the no offloading condition and they perceived a significantly lower effort. However, we did not specifically manipulate cognitive effort or related variables, and some of our drill-down analyses were post-hoc, so the support we provided is only preliminary and indirect.

Considering the marked differences between our spatial planning task and the simpler tasks previously employed in the cognitive offloading literature, the convergence of our findings with previous ones is very encouraging. Indeed, our task required a combination of skills ranging from visual perception to cognitive planning, and it allowed participants to spontaneously employ more types of offloading strategies in combination, which was not the case in previous studies. Thus, it was not a given that previous results on cognitive offloading would be observed also in our route planning domain.

However, one of our predictions based on previous offloading works was not supported. Not only did we not observe the postulated negative relationship between cognitive skills and tendency to offload cognition (Gilbert, 2015b; Gilbert et al., 2020; Meyerhoff et al., 2021), but our results indicated a significant positive relationship. Therefore, we performed an exploratory mediation analysis to delve deeper into this unanticipated result. Our findings revealed that participants with better visual search ability used more offloading strategies, and this supported their performance in the high difficulty maps. Our explanation for these unexpected findings is related to the possibility that participants in the offloading *vs*. no offloading conditions used different strategies to complete the route planning task (see Hayes-Roth, 1979; Hayes-Roth & Thorndyke, 1979; see also Wiener, Ehbauer, & Mallot, 2009). Indeed, the offloading strategies employed by participants may have transformed the planning task into a more visual task, if compared to the task faced by participants in the no offloading condition. Marking out all the locations to be visited and the ordering constraints would allow participants to transform a planning task into a sort of travelling salesman problem, which can be solved very effectively by using visual strategies (MacGregor, & Chu, 2011; MacGregor & Omerod, 1996). Research on the travelling salesman problem also showed that individual differences in visuo-spatial skills are positively associated with performance (Burns et al., 2006). In line with this interpretation, participants with better visual search skills may have strived more to transform the route planning task into a visuo-spatial task via offloading in order to solve the task through the use of their more effective visuo-spatial skills. According to this explanation, participants in the no offloading condition, not being allowed to make the transformation of route planning into a visual task, had to resort to an incremental planning strategy (Hayes-Roth & Hayes-Roth, 1979; Fum & Del Missier, 2000a; Reitter & Lebiere, 2010; see also section 1.2), mentally planning partial paths and memorizing the letters of the corresponding locations in order to report them later. The greater involvement of working memory in the no offloading condition is supported by the observation of a positive relation between working memory skills and planning performance in the more difficult maps.

Beyond the selective patterns of correlations, this explanation is supported by another piece of evidence. Indeed, in the high difficulty maps we observed a significant difference between participants in the offloading *vs*. no offloading conditions in the time needed to write the planned solution within the appropriate frame on the map sheet, despite no difference in the overall completion time of the task.

In particular, participants in the no offloading condition took much more time to write the planned route than participants in the offloading condition in the high difficulty maps (offloading condition × map difficulty interaction, $F(1,102) = 24.89$, $p < .001$, $\eta^2_p = .20$; post hoc comparison between offloading vs. no offloading condition in high difficulty maps, $t(102) = 4.80$, $p < .001$ ($M_{no\_off} = 46.08$, $SD_{no\_off} = 33.17$; $M_{off} = 22.96$, $SD_{off} = 13.65$). This may indicate that participants in the no offloading condition planned incrementally, reporting chunks of route temporarily kept in working memory, while participants in the offloading conditions planned the path visually, supported by the offloading strategies they applied, and reported the final route only after the conclusion of this procedure.

These findings point to the need to consider with greater attention the strategic variability associated with cognitive offloading in the accomplishment of cognitive tasks, as offloading may change the strategies employed, and this is more likely to emerge in complex tasks where a variety of different strategies, relying on different mixtures of cognitive skills, can be deployed (e.g., Simon, 1975). While a negative relation between cognitive abilities and the tendency to offload cognition can be observed when offloading allows participants with lower ability to be externally supported (Meyerhoff et al., 2021), the same relation can become positive when offloading allows participants with better ability to employ their skills more effectively (e.g., Hoffmann et al., 2014). More generally, our findings suggest that a complex spatial planning task can be carried out in different ways and requires different mixtures of cognitive skills when the participants are allowed *vs.* not allowed to offload cognition. It could be hypothesized that offloading is not only about reduction of cognitive load or effort (in a unitary and quantitative sense), but also about changing (re-designing) the task so that its cognitive demands are qualitatively changed (Norman, 1991).

One additional finding that may deserve further investigation is that participants who completed the high difficulty maps first perceived the task as less difficult and less effortful to be completed. From the additional analyses performed (see note 6), it seems that these findings cannot be explained by the higher rate of offloading in the low difficulty maps by participants in the offloading condition who completed the high difficulty maps first. An alternative possibility could involve an anchoring effect.

Indeed, participants in all the conditions may have anchored to the perceived difficulty and effort of the first block of maps when providing their judgments on the second block of maps.

Concerning the domain-specific contribution of our study, we have highlighted four specific offloading strategies spontaneously used in route planning on maps, often employed in combination, and showed that their implementation contributes to a better performance and to the reduction of perceived cognitive effort. We also showed that route planning with or without the opportunity to offload cognition on a map requires different skills. This points to the need of integrating theories of route planning with a much more explicit account of the way in which externalizing cognition on a map changes the planning process by modifying the transactions between the individual and the environment (Hutchins, 1995; Norman, 1991; Zhang, 1997).

2.4.2 Applied Implications

The results of our study suggest some potential applied implications. First, our study offers hints on how the externalization of cognition could help human route planning on maps. There are some real-world situations in which navigation systems cannot be used or are unlikely to be used (e.g., plan a route within a park, a mall, or any other relatively small or non-digitized area). In these situations, the offloading strategies we identified could be useful to plan a path with a real or digital map.

Second, the four offloading strategies we identified could be incorporated in the design of training or rehabilitation programs of neuropsychological patients with deficits in spatial planning. For example, individuals who experience difficulties in spatial planning due to discontinuity in their planning efforts (e.g., Basso et al., 2001) could be taught to keep an external track of their plan as they develop it, as this may support their monitoring process. Training or rehabilitation programs may benefit also from the general approach we have employed to discover effective offloading strategies: observing which strategies participants use when allowed to offload cognition and evaluating their effectiveness. Indeed, the same approach could be applied in other complex tasks and domains.

Third, the strategies we have identified may also be incorporated into systems supporting interactive route planning to improve path planning and route visualization. Indeed, digital maps lend themselves to the implementation of some of the offloading strategies we identified through our paper-based task, such as marking out multiple locations of interest or visualizing possible routes through all of them. However, implementing these strategies in navigation systems, particularly with multiple locations and constraints, is currently rather cumbersome (e.g., Google Maps), showing room for improvement.[9] Some of the insights we gained can be used for informing such improvements. For instance, our findings suggest that the possibility to easily specify ordering constraints between locations directly on the digital map and having them externally represented could improve actual interfaces.

Fourth, the offloading strategies we observed could also be used to support the human-centered design of less automated route planning systems actively promoting cognitive map learning, considering the potential cognitive drawbacks of fully automated systems (Brügger et al., 2019). For instance, the user may be asked to select the locations to be visited directly on the digital map and specify the constraints between locations in a way more similar to what is naturally done on a paper map (e.g., tracking a directed arrow). Then the system may suggest a route respecting the constraints and showing the path on a survey view of the map (e.g., as in Google Maps; Shelton & Gabrieli, 2002). Next, the user may be asked to review the path and change it interactively if needed to get a better or preferred solution. These proposed changes may be tracked directly on the map, as happens in a paper map. Finally, the user can preview the path in a route view, seeing each target location in the planned order (such as in Street View; Shelton & Gabrieli, 2002), but additionally being prompted to pay attention to each target location to foster a deeper encoding.

[9]Popular free technological platforms like Google Maps typically provide the shortest route between pairs of locations. When considering multiple locations and ordering constraints between locations, they compute the route using the user-provided sequence of locations, rather than providing a sequence of locations that minimizes the whole length of the route. This implies that the use of navigation systems does not eliminate the need for planning and does not necessarily lead to the full automation of this process, especially when the planning task becomes more complex.

2.4.3 Limitations and Future Research Directions

Our study has some limitations that can be addressed by future research. First, our explanation of the unexpected finding of a positive relation between visual search ability and use of offloading strategies in difficult maps, which refers to the use of a visual strategy, needs to be further corroborated. Future studies may employ a variety of methods to this aim, from dual task, to process tracing, to the manipulation of critical features of the maps or of the task.

Second, further research could also employ the choice/no choice paradigm to deepen the investigation of the costs and benefits of the offloading strategies we have identified (see Gilbert et al., 2023). Moreover, map difficulty could be increased or manipulated in different ways, to further understand the determinants of the use of offloading strategies and of task performance.

Third, in addition to the use of a relative evaluation of planning performance based on the best route identified in the sample of participants, an absolute evaluation based on an algorithmically determined shortest route could be used, provided that a computationally effective algorithm could be designed for our kinds of problems. However, this will not affect the results and the conclusion of our study, as they are based on relative comparisons of performance.

Fourth, different kinds of task designs or maps could be used to appraise the generalizability of our findings (e.g., maps with landmarks, semantically rich names or pictures, non-linear paths; tasks with time or cost constraints, priorities about locations to be visited).[10]

Fifth, it would be worth investigating different populations, such as children, older adults, or individuals with planning deficits, who may benefit even more from the use of offloading strategies on maps.

[10] Our study has not examined the role of map-reading skills in the context of route planning, considering the abstract nature of our maps. According to a recent review by Arthurs et al. (2021), two studies explicitly defined map reading. Bluestein and Acredolo (1979) characterized it as the capacity to discern one's position in a three-dimensional, large-scale space using information from a two- or three-dimensional small-scale representation (p. 691). Dobson (1977) described it as a process whereby sensed map data are converted into neuro-electrical activity and interpreted through perceptual schemata in relation to identification, classification, and coding. Both definitions suggest that map reading skills are related to navigation tasks within an environment rather than to route planning on maps. However, it could be useful to consider how the abilities to understand semantically rich maps and their symbols and to make related spatial inferences are associated with route planning on maps with/without the possibility to offload cognition.

Sixth, it could be interesting to understand how individuals offload cognition while interacting with digital maps or navigation software. From a methodological perspective, it should be noted that paper-and-pencil tasks, like the one we used, allow us to understand how individuals spontaneously devise a variety of offloading strategies in relatively unconstrained situations. In contrast, the use of computerized tasks typically limits such possibility, as the actions that can be done with an interface are necessarily constrained by the interface design. However, some room for human action typically remains. Moreover, changes in the interface design may have significant effects in the way a task is performed and on performance (e.g., O'Hara & Payne, 1998). Therefore, the ways in which different interface designs allow individuals to offload cognition and their consequences still need to be investigated.

All these limitations could be addressed by future research, but our study already leads to the conclusion that planning routes on maps, with or without the possibility to offload cognition, makes a significant difference, and that participants are able to deploy, spontaneously and in a partly adaptive way, a variety of offloading strategies in this kind of spatial planning. This highlights the importance of conceiving human planning in close relation with the surrounding environment and with the actions that can be performed with the tools and artifacts populating the environment, attentively considering the consequences of these actions on cognition and thinking.

STUDY 2 AND STUDY 3.

TRANSFER EFFECTS OF COGNITIVE OFFLOADING IN ROUTE PLANNING

3.1 Introduction

Several studies have investigated the consequences of cognitive offloading on episodic memory, particularly focusing on how reliance on external aids affects memory for offloaded information. For example, Sparrow et al. (2011) found that participants who believed the statements they had read were saved to a computer demonstrated poorer recall of those statements compared to participants who thought the statements had been erased. This effect occurred regardless of whether participants were explicitly instructed to remember the information. Notably, those who believed the information was externally stored were more likely to remember where it was stored (e.g., the specific computer folder) rather than the content itself, suggesting that the availability of external storage can lead individuals to prioritize remembering *where* to access information over the details of the information itself. Other studies have examined the consequences of cognitive offloading on visual episodic memory, specifically investigating how taking photos affects memory for the photographed objects. Henkel (2014) found that, after a museum tour, individuals remembered fewer details about objects they photographed compared to objects they simply observed. However, zooming in while taking photos made this difference disappear. Expanding on this, Soares and Storm (2018) demonstrated that the photo-taking impairment effect persisted even when participants knew they would not have future access to the photos, suggesting that it was the act of offloading itself, rather than reliance on future access, that disrupted memory.

Interestingly, it was also observed that searching for information online can lead individuals to overestimate their internal knowledge. Fisher et al. (2015) found that participants who had access to the Internet to search for answers to general knowledge questions, compared to participants who did not have Internet access, rated themselves as more knowledgeable on a subsequent set of unrelated questions. Importantly, this inflated sense of knowledge occurred even though participants could not search online for information on these new topics. These findings point out that Internet access may interfere with individuals' ability to accurately distinguish between internal and external sources of knowledge, ultimately leading to an inflated sense of their own knowledge (see also Ward, 2021).

Subsequent research has also highlighted potential benefits of storing information externally, specifically showing that saving a list of words to a digital file before studying a new list can improve memory for the second list (Storm & Stone, 2015). Thus, while offloading information to digital tools can impair memory for the offloaded material, it may also facilitate learning new information, in some cases possibly due to the reduction of proactive interference. In line with this, Eskritt and Ma (2014) observed that offloading information by note-taking can alleviate proactive interference in a visual memory task. In their first experiment, participants viewed an array of matched pairs of cards depicting common objects. Some participants were allowed to take notes, while others had to study the cards and their locations without aids. During the recall phase, note-takers were unexpectedly deprived of their notes, and all participants were tested on both the identity and location of the cards. Although note-takers recalled the identities comparably to those who relied solely on memory, they remembered significantly less about the locations. Eskritt and Ma hypothesized that note-takers had “intentionally forgotten” the location information through their annotation, while focusing on memorizing the identities of the cards. To test this hypothesis further, the authors conducted a second experiment in which participants, once again divided into note-taking and study groups, recalled card identities and locations over multiple trials. However, in the final trial, no notes were permitted. Note-takers showed better performance across all trials, suggesting that note-taking reduced proactive interference from previous trials both when notes were being taken, and in the final task where only internal memory could be used. These findings align with the authors' hypothesis: by offloading information through note taking, participants “intentionally forgot” some of the information, which ultimately protected them from interference across trials.

Building on these findings, Finkler and Meyerhoff (2024) further investigated whether the protective effect of note taking stemmed from reduced proactive interference or from a decrease in mental fatigue. In their first experiment, they replicated the protective effect of note taking: participants who took notes during initial trials of a card-matching task, compared to those who did not, demonstrated better memory performance in a final, unaided trial. Importantly, in their second experiment, they additionally found that this beneficial effect persisted when participants were tested on repeated stimuli only, and not on novel stimuli. This suggests that note taking specifically enhanced memory by reducing proactive interference from previously encountered items, rather than by reducing mental fatigue or broadly improving subsequent memory.

Cognitive offloading was also found to affect spatial memory, particularly of navigated environments. Fenech et al. (2010) showed that using acoustic turn-by-turn navigation impaired participants' ability to recognize scenes from a route they had navigated, suggesting that reliance on navigational aids reduces environmental encoding, eventually impairing spatial knowledge. Gardony et al. (2015) further supported these findings, demonstrating that both tonal and navigational aids, although improving navigational efficiency, significantly impaired participants' ability to remember landmarks and form a mental map of the environment. The authors argued that this occurred because navigational aids distracted navigators during environment encoding, leading to an impoverished and poorly integrated spatial mental representation.

Overall, these studies provide important insights into the consequences of cognitive offloading on verbal, visual, and spatial episodic memory. Across different modalities, the general pattern of findings suggests that, although cognitive offloading can enhance immediate task performance or free cognitive resources for other purposes, it often impairs memory for the offloaded information. However, it may also facilitate performance in further memory tasks due to reduced proactive interference. Thus, offloading cognition can lead to a complex performance pattern, with both positive and negative consequence in the short term and in the longer term.

However, to the best of our knowledge, beyond Eskritt and Ma (2014) and related research, only one study has directly investigated the consequences of prior versus non-prior use of cognitive offloading on subsequent performance of the same task. Scarampi and Gilbert (2020) conducted three experiments that focused on memory for delayed intentions, specifically exploring how prior use of reminders affects subsequent performance. In Experiment 1, participants were either forced to use external reminders or to rely solely on their memory during the first phase of a delayed intention task. The results showed that prior use of reminders, compared to non-use, had little to no detrimental effect on performance in a subsequent task phase where reminders were no longer allowed. In Experiment 2, participants who were initially forced to use reminders were more likely to continue using them in a later phase where they could freely choose between setting external reminders or relying on internal memory. This contrasted with participants who had initially relied on memory alone, suggesting a perseverative use of the previously experienced strategy. In Experiment 3, Scarampi and Gilbert found that performing just one trial with the opposite strategy (i.e., using reminders after relying solely on memory, or vice versa) eliminated the strategy perseveration effect in the subsequent free-choice phase of the task.

Outside the context of offloading research, the first empirical evidence for the perseverative use of a practiced or learned strategy was provided by the seminal work of Luchins and Luchins (1950). Through a series of studies on transfer effects in problem-solving tasks (water jar problems), the authors consistently found that, once participants became accustomed to a specific strategy, they struggled to abandon that strategy, often continuing to use it even when it was no longer efficient or necessary. This pioneering work introduced the concept of the "Einstellung effect", which occurs when individuals persist in using a familiar strategy, even when a simpler or more efficient approach is available, ultimately limiting their flexibility in problem-solving (for more recent works on Einstellung effect, see e.g. Bilalić & McLeod, 2014; Bilalić et al., 2008; Ohlsson, 2011). Part of the results from Study 1 could be also read as a sort of Einstellung effect. Indeed, we observed that participants in the offloading condition, after starting with high difficulty maps, continued to use approximately the same number of offloading strategies when moving to low-difficulty maps. However, this approach did not confer them any advantage but, instead, slowed them down compared to participants in the no-offloading condition. However, in our case, a concurring or alternative explanation is represented by the reduction in cognitive effort, which has been proposed as a main driver of offloading (Gilbert et al., 2023; Sachdeva & Gilbert, 2020). Early work by Luchins and Luchins and subsequent research highlights the enduring impact of prior strategy use on subsequent cognitive tasks, a theme that resonates with the limited existing research on cognitive offloading dealing with this issue (e.g., Scarampi & Gilbert, 2020). This research also suggests the need to consider more carefully whether and how different initial strategies might influence subsequent task performance when a new strategy is required.

Important insights into how initial strategies can shape adaptation to a new strategy have been provided by O'Hara and Payne (1998). They conducted a series of studies examining whether manipulating the cost of implementing actions during problem-solving affected participants' level of planning. In their first two experiments, participants solved a series of 8-puzzle tasks. Some of them solved the puzzles in a high-cost condition, where each move involved a lengthy command, and the others solved them in a low-cost condition, which involved simple key presses. Results showed that participants in the high-cost condition used fewer moves, indicating more planful and deliberate actions, while those in the low-cost condition relied more on immediate, display-based actions. In their third experiment, O'Hara and Payne explored how these initial strategies would transfer when all participants shifted to a different low-cost condition. They found that participants initially trained in the high-cost condition performed more efficiently, both in terms of time and moves, compared to those who had initially trained in the low-cost condition. This suggested that experience with a more planful strategy promoted better adaptation, by fostering better learning of the task to be performed. In their final experiment, the authors tested whether this advantage transferred to a different problem domain, finding no benefit, which indicated that the planning benefits were task specific. O'Hara and Payne's findings underscore the impact of initial strategy on the shift to a different strategy, in particular suggesting that, when a strategy requires more planning, it can lead to a deeper understanding of the task's underlying structure, thus facilitating an easier adjustment to new strategies when these can be fostered by this understanding and the task remains the same.

Building on evidence from previous studies, Study 2 and Study 3 aimed to extend the investigation to the consequences of cognitive offloading. While earlier research on the consequences of cognitive offloading primarily focused on episodic or prospective memory with relatively simple tasks, our research employed a more complex task in the domain of route planning.

Notably, past studies have shown how a strategy initially practiced influences subsequent task performance when that strategy is either no longer available (Eskritt & Ma, 2014; Scarampi & Gilbert, 2020), less efficient (Luchins & Luchins, 1950), or when the implementation cost of its operators is changed (O'Hara & Payne, 1998). However, in the context of cognitive offloading, no study has yet considered the bidirectional nature of transfer effects, which means examining the consequences of a certain initial strategy on the shift to a different strategy, while simultaneously investigating the shift in the opposite direction.

In Study 2 and Study 3, we tried to fill this gap by examining how allowing/not allowing the opportunity to use offloading strategies affected subsequent performance when this opportunity was either removed or introduced. In Study 2, participants were free to spontaneously devise their own offloading strategies while performing a route planning task. In Study 3, the strategies spontaneously devised by participants in Study 2 (which aligned with those devised by participants in Study 1) were supported by specific external visual aids.

Beyond broadening the investigation of cognitive offloading into a new domain, Studies 2 and 3 aimed to provide deeper insights into how the use of external visual aids shapes cognitive strategies underlying route planning, particularly regarding the ability to shift between reliance on external supports and internal cognitive resources. Drawing on insights from our previous study, we expected that the introduction or removal of the opportunity to use offloading strategies would have differential impacts on task performance, resulting in asymmetric transfer effects (see the next section for our hypotheses and their rationale).

3.2 Study 2

3.2.1 Research Aims and Hypotheses

The main aim of Study 2 was to investigate whether learning a route planning task with or without the possibility to offload cognition would affect the subsequent performance of the same task when the opportunity to offload cognition was removed or introduced. To this aim, participants were first presented with a training block of the route planning task (see Study 1) and required to complete it under different between-subjects conditions. A subset of participants was trained with the opportunity to offload cognition by freely using a pen (as in the offloading condition in Study 1), whereas other participants were trained without such opportunity (as in the no offloading condition in Study 1) or asked to keep their hands still while planning. The latter condition prevented participants from offloading cognition onto the body by using their hands when external offloading via pen use was precluded (Risko & Gilbert, 2016).

After the training block, participants completed two test blocks. In one block, participants performed the route planning task under the same conditions as in their training. In the other test block, participants who were trained with pen use were prevented from using it or asked to keep their hands still, while participants who were trained without the pen were allowed to use it.

We hypothesized that asymmetric transfer effects would emerge on route planning performance. Specifically, in accordance with prior research indicating that cognitive offloading typically enhances task performance (e.g., Gilbert, 2015a), as well as with findings from our previous study, we expected that removing the opportunity to offload cognition would impair route planning and that introducing such opportunity would improve it. Crucially, we further expected that the decline in performance following the removal of offloading would be greater than the improvement following its introduction (H1). This expectation was grounded in evidence from our previous study. We found that route planning performance was correlated with visual search ability in the offloading condition, whereas it was correlated with spatial working memory in the no offloading condition. These findings, together with the analysis of solution writing times, suggested that participants who had the opportunity to offload cognition used a planning strategy that was mainly visual, similar to those intuitively used when performing traveling salesman problems (see e.g., MacGregor, & Chu, 2011; MacGregor & Omerod, 1996). In contrast, for participants who were not allowed to offload cognition, our findings suggested that they relied mostly on an incremental and more demanding working memory-based strategy. We expected that switching from a spontaneous, primarily visual strategy to a more taxing, working memory-based strategy would be more challenging than switching in the reverse direction. Indeed, picking up the intuitive visual strategy should be easy when offloading cognition into the map by marking it allows to transform the planning task in a sort of traveling salesman problem (see Study 1). On the contrary, having used such a visual strategy in the past can get into the way when it cannot be used anymore, and the participants need to switch to a more thoughtful and demanding incremental planning one.

It is interesting to note that this hypothesis is different from a simple einstellung-based one (Bilalić & McLeod, 2014; Luchins & Luchins, 1950) and different from the one that would stem from the work on the implementation cost of the operators (O'Hara & Payne, 1998). Indeed, according to einstellung-based research, previous experience with a different strategy should always get into the way when a new strategy needs to be devised, showing negative transfer effects on subsequent performance. On the other hand, the research on the implementation cost of the operators suggests that a more planful strategy, fostered by a higher implementation cost of operators, promotes a better learning of the structure of task, with positive transfer effects on subsequent performance. But these hypotheses do not refer to situations in which strategy changes occur because the opportunity to offload cognition is manipulated, and this can lead to deep changes in the nature of the task to be performed, even though it is apparently the same.

Indeed, in our case, we hypothesized that the degree of positive/negative change in performance depends on the specific strategy shifts experienced by participants as a consequence of the introduction/removal of the opportunity to offload cognition. This implies the need to carry out a finer analysis of the strategies employed in the different conditions and of the way in which the previous use of a strategy (with/without the opportunity to offload cognition) may affect the shift to a different one (without/with the opportunity to offload cognition), with positive or negative consequences on performance.

In addition to the effects on performance, we further hypothesized that changing the availability of cognitive offloading would also change participants' perceptions of the task. This hypothesis was also based on evidence from our previous study, which showed that participants who were given the opportunity to offload cognition perceived the task as less difficult and less effortful than those who were not given such opportunity. As a consequence, in the present study, we expected that removing the option to offload cognition would lead to an increase in perceived task difficulty and effort, whereas introducing this option would lead to a decrease in these perceptions (H2).

A second aim of the study was to examine whether being trained on our route planning task with *vs*. without the option of offloading cognition via pen use would lead to different degrees of offloading strategy adoption in the subsequent test phase where pen use was allowed. All participants completed one test block with the free use of the pen, but their experience with this option varied according to their training conditions (see section 3.2.2.3). We expected that participants previously trained with the free use of the pen would make greater use of offloading strategies in the test block with free pen use than those previously trained without the use of the pen (H3). This expectation is consistent with the evidence indicating that prior experience with offloading strategies promotes their subsequent use (Scarampi & Gilbert, 2020; see also Study 1).

3.2.2. Method

3.2.2.1 Participants

An a-priori power analysis indicated a minimum sample size of 136 participants to detect a medium effect size for our experimental design ($f = 0.25$, $\alpha = 0.05$, $1 - \beta = 0.80$). We recruited a total of 142 undergraduates for this study. Participants who did not adhere to the route planning task instructions (n = 4), and those whose performance scores deviated by more than three standard deviations from the sample mean on at least three measures of the route planning task (n = 1) were excluded from the analyses. Our final sample consisted of 137 participants (84.67% female, age: M = 22.36, SD = 5.78). The study was approved by the Ethical Committee of the University of Trieste, and informed consent was obtained from each participant at the beginning of the experimental session.

3.2.2.2 Experimental Design

All participants underwent individual experimental sessions, in which they were presented with a training block followed by two test blocks of the route planning task. The training condition was manipulated between participants, and one of the test blocks was performed under a different condition from the training.

During the training block, two groups of participants were allowed by the instructions to freely use a pen on the map sheets during the task, as in the offloading condition of Study 1. Another group was instructed to use the pen only to indicate the final routes, as in the no offloading condition of Study 1. A fourth group was additionally instructed to keep the hands still on the table until the final routes were fully identified and ready to be reported. This condition was included to ensure that participants did not engage in cognitive offloading by using the hands, a possibility that remains when pen usage is restricted but no specific instructions are provided regarding hand movements.

After completing the training block, participants proceeded to two test blocks. In one test block, participants performed the task under the same condition as in the training block (training-consistent block). In the other block (training-inconsistent block), both groups trained with the pen were now restricted from using it, with one group also being restricted from using their hands. Conversely, the groups initially not allowed to use the pen (or both the pen and hands) were now allowed to use the pen.

This setup resulted in a 4 (group: pen→no-pen *vs*. pen→no-hand *vs*. no-pen→pen *vs*. no-hand→pen) × 2 (test block: training-consistent *vs*. training-inconsistent) mixed experimental design.

3.2.2.3 Procedure

At the beginning of the experimental sessions, participants read and accepted the informed consent form and were then randomly assigned to one of four groups:

1. pen→no-pen group, with pen use allowed during the training block, and preclusion of pen use during the training-inconsistent test block (n = 35);

2. pen→no-hand group, with pen use allowed during the training block, and preclusion of both pen and hand use during the training-inconsistent test block (n = 36);

3. no-pen→pen group, with preclusion of pen use during the training block, and pen use allowed during the training-inconsistent test block (n = 33);

4. no-hand→pen group, with preclusion of both pen and hand use during the training block, and pen use allowed during the training-inconsistent test block (n = 33).

Participants read the instructions for the training block of the route planning task, completed a practice map, and estimated how much they thought the routes they were about to plan would deviate from the shortest possible routes in the upcoming block. After completing the maps for the training block, participants reported how much they believed they had actually deviated from the shortest routes, as well as how difficult and effortful they found the completion of the training block. Participants were then given a brief filler questionnaire (Woo et al., 2014) to introduce a short, nondemanding delay between the training block and the test blocks. Once they completed the questionnaire, participants proceeded to the test blocks, which were presented in a counterbalanced order. The instructions for the training-consistent block informed participants that they had to perform the route planning task exactly as they had practiced it during the training stage. The instructions for the training-inconsistent block informed participants that they had to perform the route planning task under a different condition from that practiced during the training, with specific differences explained according to their group. Both test blocks were delivered following a procedure like the one of the training block, consisting of instructions, performance prediction, route planning task, retrospective judgment of performance, and evaluation of perceived difficulty and effort.

Finally, participants completed a computerized visual search task (MFTC; Marra et al., 2013), answered demographic questions, and performed the Symmetry Span Task (Kane et al., 2004). The total duration of the procedure was approximately 70 to 80 minutes. A debriefing session was held after data collection, during which the full aims of the study and further details were disclosed to participants.

3.2.2.4 Materials

3.2.2.4.1 Route Planning Task

The route planning task is a paper-and-pencil task that presents participants with a series of maps of abstract locations. Participants were required to identify the shortest possible route on each map from a specified starting place to a designated final place, visiting specific intermediate locations and adhering to prescribed ordering rules (see Study 1). While the task was as in Study 1, some changes were made to the procedure to better address the objectives of the current study.

First, while the previous study included both high- and low-difficulty maps, the current study used only high-difficulty maps to increase task sensitivity. This decision was based on the observation that performance differences between conditions allowing *vs*. precluding cognitive offloading (using a pen) were particularly pronounced on high-difficulty maps.

Second, whereas participants in Study 1 completed two blocks of eight maps each, the current study employed a different structure: a training block of six maps followed by two test blocks of three maps each. This design was chosen to examine how training with *vs*. without the option to offload cognition would influence subsequent performance when the option was either removed or introduced. Given that low-difficulty maps were excluded, the total number of maps was reduced from 16 to 12 to avoid excessively long sessions and potential participant's fatigue.

The maps were selected from those used in Study 1, specifically choosing the 12 most comparable in terms of speed and accuracy. The six maps assigned to the training block were overall comparable to the six maps used in the test blocks. Each map presented a starting point, an ending point, eight intermediate locations to be visited, two ordering rules (each constraining the order of visit of a pair of intermediate locations), and ten distractor locations. To control for order effects, the presentation order of maps was randomized within the training block and across the two test blocks.

Participants were provided with a booklet containing a series of layered sheets, each displaying a map along with its specific directions on the right (e.g., start and end places, locations to be visited, and ordering rules to follow). Participants were instructed to plan the shortest route as quickly and accurately as possible and, only after fully identifying their solution, to report it in a designated response frame located beneath the map. This required them to write down the locations (letters) in the exact order as they planned to visit them.

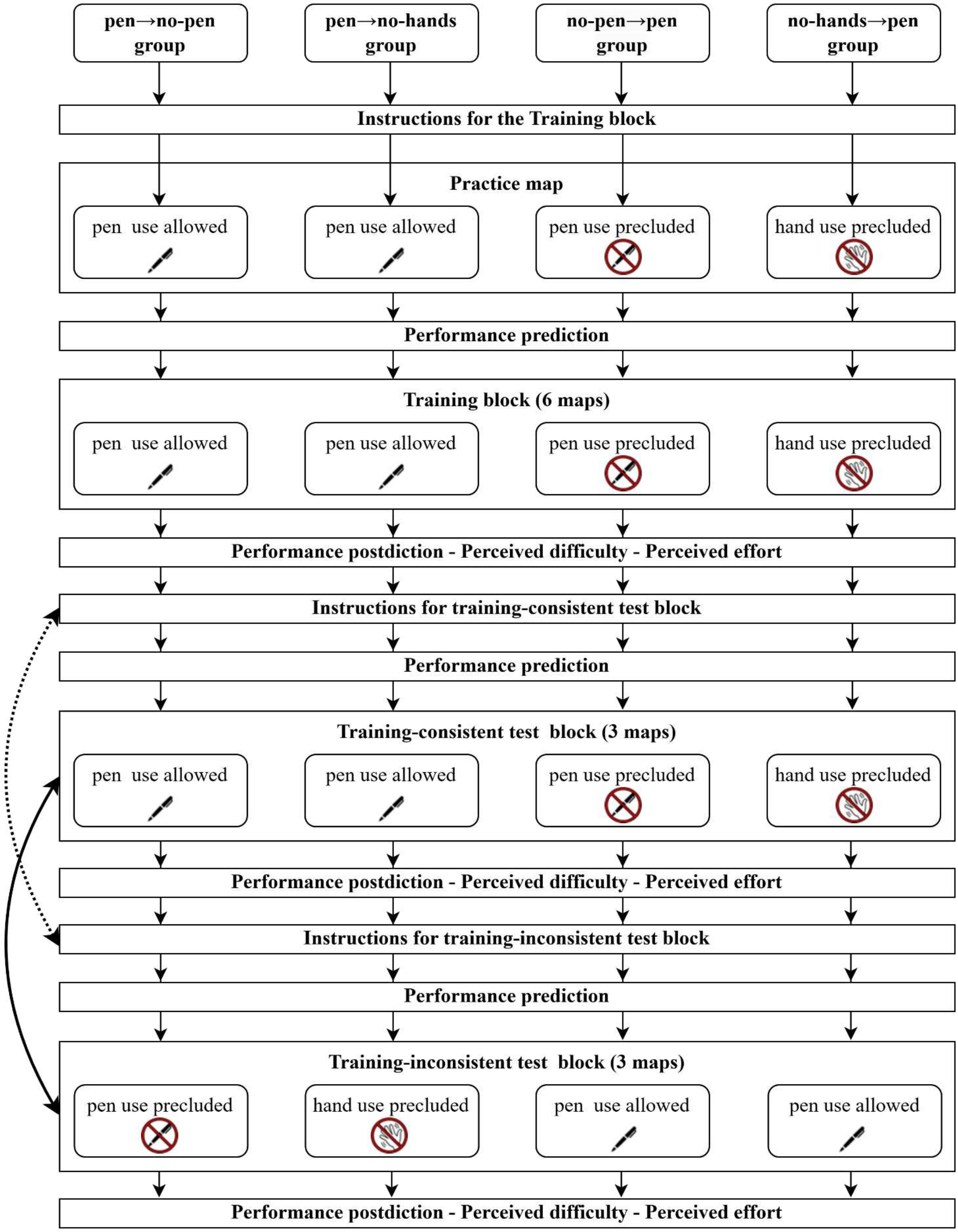
The instructions for the training block were identical across all groups, with the following exceptions. Participants in the pen→no-pen group and the pen→no-hand group were informed that they could use the pen on the map sheet at any point during the task, even before recording their final route. Participants in the no-pen→pen group were instructed to use the pen exclusively to record the final route in the designated response area, with no instructions regarding hand use. Participants in the no-hand→pen group were also instructed to use the pen exclusively for recording the final route in the designated response area, but were additionally required to keep their hands still on two blue plastic rectangles attached to the table on the two sides of the booklet until they had fully planned their route and were ready to record it.

For the training-consistent test block, participants were informed that they would be presented with additional maps and were required to complete them in the same manner as during the training block. For the training-inconsistent test block, participants were also informed that they would encounter new maps and that they were still required to identify the shortest possible routes as quickly and accurately as possible, but with changes in pen and, in some cases, hand usage. Participants in the pen→no-pen group were instructed to use the pen only to write down their final route after fully identifying it, mirroring the training of the no-pen→pen group. Participants in the pen→no-hand group were additionally required to keep their hands still on the blue plastic rectangles, like the no-hand→pen group during the training. Participants in the no-pen→pen and no-hand→pen groups were informed that they could now use the pen freely during the task, as was allowed for the pen→no-pen and pen→no-hand groups during training. Thus, participants who had been allowed to use the pen during the training block were no longer allowed to use it in the training-inconsistent block, but could use it in the training-consistent block. Conversely, participants who had been restricted from using the pen during the training were allowed to use it in the training-inconsistent block but not in the training-consistent block. Notably, all participants completed one test block allowing pen use, and one without pen use. Given that Scarampi and Gilbert (2020) observed that perseveration effects from the training phase to the testing phase can be eliminated after just one trial using the opposite strategy, we counterbalanced the order of test blocks to control for this possibility and block order effects. For each map, once recorded the entire route, participants were required to flip the sheet and place it in a box on their right before moving on to the next map (refer to Figure 7 for an overview of the procedure followed in administering the Route Planning task).

The scoring procedure followed that used in Study 1. First, we calculated the proportion of maps completed without errors (i.e., correct start/arrival locations, inclusion of all intermediate locations, and adherence to ordering rules). For these error-free maps, we assessed planning accuracy by identifying the shortest possible route and calculating the proportional deviation above this route using the formula: *proportional deviation* = (actual route length - shortest route length) / shortest route length. Higher values indicated greater deviation from the optimal route. For planning speed, we recorded the time taken to complete each error-free map, measured as the interval between the flipping of consecutive map sheets.

Figure 7

Overview of the Route Planning Task Administration Procedure in Study 2



Note. Bidirectional arrows indicate the counterbalanced presentation of test blocks and associated instructions.

3.2.2.4.2 Coding and scoring of offloading strategies

All maps completed under the condition allowing pen use (i.e., maps from the training block and training-consistent test block for the pen→no-pen and pen→no-hand groups, and maps from the training-inconsistent test block for the no-pen→pen and no-hand→pen groups) were evaluated by an expert rater. The offloading strategies employed by participants were categorized using the schema from our previous study (see Study 1). This schema identifies four possible categories of offloading strategies: (1) marking starting and arrival locations; (2) marking intermediate locations to be visited; (3) marking ordering constraints; and (4) keeping track of the plan. As in Study 1, no additional offloading strategies were identified in the present study.

Maps were individually assessed for the presence of each offloading strategy using a dichotomous measure: 0 for non-use and 1 for use. Each map was also evaluated for the total number of offloading strategies implemented, with scores ranging from 0 (no strategies used) to 4 (all strategies used). The number of strategies employed was then averaged across maps belonging to the same block (i.e., training block and training-consistent test block for the pen→no-pen and pen→no-hand groups, and training-inconsistent test block for the no-pen→pen and no-hand→pen groups).

3.2.2.4.3 Metacognitive Performance Judgments, Perceived Difficulty, and Effort

Performance Prediction and Postdiction. After reading the instructions for the training block and completing the corresponding practice map, participants were asked to estimate how much they would deviate from the shortest possible routes on the upcoming maps. They responded on a 10-point scale, ranging from *I will deviate a lot* to *I will deviate very little*. Once they finished the training maps, participants were asked to retrospectively estimate how much they had actually deviated from the shortest routes, using the same scale (*I deviated a lot – I deviated very little*). This procedure (Gilbert, 2015) was then repeated for both test blocks (i.e., the training-consistent block and the training-inconsistent block). However, in the test blocks, no practice maps were provided before asking participants to predict their performance, as they were already familiar with the task.

Perceived difficulty and effort. After finishing each block of maps, participants were also asked to rate the difficulty of the maps they had just completed and the amount of effort they had put into completing them, using a 7-point scale ranging from *very little* to *very much*.

3.2.2.4.4 Measures of Cognitive Abilities

Visual Search. To evaluate visual search ability, we administered a computerized version of the Multiple Features Target Cancellation task (MFTC; Marra et al., 2013). Participants viewed a grid of 80 small squares, each containing two differently oriented lines. After fixating on a probe (a small square with two oriented lines) positioned above the grid, they were required to identify as quickly and accurately as possible the squares that matched the probe by clicking on them. Timing began when participants entered the screen displaying the targets (n = 13) and distractors (n = 67) and ended when they pressed a bottom-right arrow to confirm completion. Accuracy was calculated using the following formula: [(correct responses / 13) + ((67 - false alarms) / 67)] / 2 (Marra et al., 2013).

Spatial working memory. To assess spatial working memory (WM) capacity, we employed the Symmetry Span task (Kane et al., 2004) from the Attention & Working Memory Lab (https://englelab.gatech.edu/). In each trial, participants viewed a series of 8 × 8 matrices with black-and-white squares and determined whether each matrix was symmetrical along its vertical axis (process component). Following each symmetry judgment, a red square was briefly displayed within a 4 × 4 matrix (storage component). Each sequence comprised 2 to 5 matrices, alternating with the same number of red squares, and concluded with a 4 × 4 matrix where participants were required to click on the squares corresponding to the previously shown red squares according to their order of appearance. Before the actual trials, participants completed separate practice sessions for the process component, the storage component, and the full task. During the actual trials, if the time spent on a symmetry judgment exceeded the average practice time by more than 2.5 standard deviations, the program automatically presented the subsequent red square and recorded the symmetry judgment as an error.

Performance was assessed using the partial score, which considers all correctly recalled red squares (Oswald et al., 2015; Redick et al., 2012).

3.2.3 Results

3.2.3.1 Preliminary Analyses

For each performance variable of the route planning task, data were trimmed within the groups of participants and blocks of maps. Values that were 3 *SD*s above or below the mean were replaced by the mean ± 3 *SD*s (for a similar procedure, see Study 1; Del Missier et al., 2021; Miyake et al., 2000). One participant was excluded from the subsequent analyses because their performance deviated by more than 3 *SD*s from the mean on all the measures in one test block. In the case of the individual-differences tasks, the replacement was conducted in accordance with the aforementioned criterion across the entire sample. Table 3 presents the descriptive statistics for all measures of the route planning task by groups and blocks.[11]

Groups did not differ significantly on gender, $\chi^2(3) = 5.40, p = .144$, visual search skills, $F(3,133) = 2.38, p = .072$, and spatial working memory, $F(3,132) = 1.38, p = .251$.

Given that the primary aim of this study was to examine how training in the route planning task under different conditions influenced subsequent performance in two test blocks where the initial

[11] Data trimming resulted in the following percentages of value substitution. For the proportion of maps completed without errors: pen→no-pen group, 2.86% for the training block, 0.00% for the training-consistent test block, and 2.86% for the training-inconsistent test block; pen→ no-hand group, 2.78% for the training block, 2.78% for the training-consistent test block, and 0.00% for the training-inconsistent test block; no-pen→pen group, 3.03% for the training block, 3.03% for the training-consistent test block, and 0.00% for the training-inconsistent test block; no-hand→pen group, 3.03% for the training block, 3.03% for the training-consistent test block, and 0.00% for the training-inconsistent test block. For proportional deviation above the shortest route: pen→no-pen group, 2.86% for the training block, 2.86% for the training-consistent test block, and 2.86% for the training-different test block; pen→no-hand group, 2.78% for the training block, 2.78% for the training-consistent test block, and 0.00% for the training-inconsistent test block; no-pen→pen group, 3.03% for the training block, 0.00% for the training-consistent test block, and 3.03% for the training-inconsistent test block; no-hand→pen group, 3.03% for the training block, 0.00% for the training-consistent test block, and 3.03% for the training-inconsistent test block. For map completion time: pen→no-pen group, 0.00% for the training block, 0.00% for the training-consistent test block, and 2.86% for the training-inconsistent test block; pen→no-hand group, 0.00% for the training block, 0.00% for the training-consistent test block, and 0.00% for the training-inconsistent test block; no-pen→pen group, 3.03% for the training block, 0.00% for the training-consistent test block, and 3.03% for the training-inconsistent test block; no-hand→pen group, 0.00% for the training block, 0.00% for the training-consistent test block, and 0.00% for the training-inconsistent test block. For visual search (MFTC accuracy scores), the trimming led to 0.01% substitutions. For spatial working memory (Symmetry Span partial scores), there were no substitutions (0.00%). Descriptive statistics for the MFTC accuracy score were as follows: $M = 0.92$, $SD = 0.09$. Descriptive statistics for the Symmetry Span task were as follows: $M = 26.60$, $SD = 7.63$.

condition was either maintained or changed, the following analyses will focus on these test blocks. Analyses for the training block are presented in Appendix B.

Table 3

Descriptive Statistics of Route Planning Performance Measures

	Proportion of Maps Without Errors			**Map Completion Time**			**Proportional Deviation Above the Shortest Route**		
Block	Training M (SD)	Consistent Test M (SD)	Inconsistent Test M (SD)	Training M (SD)	Consistent Test M (SD)	Inconsistent Test M (SD)	Training M (SD)	Consiste nt Test M (SD)	Inconsistent Test M (SD)
Group									
pen→no-pen	0.95 (0.09)	0.95 (0.12)	0.86 (0.22)	188.17 (67.31)	162.34 (51.12)	237.68 (66.70)	0.12 (0.06)	0.09 (0.08)	0.16 (0.13)
pen→no-hand	0.92 (0.12)	0.93 (0.15)	0.88 (0.16)	197.59 (55.35)	164.49 (51.78)	231.42 (85.73)	0.10 (0.05)	0.11 (0.09)	0.18 (0.14)
no-pen→pen	0.92 (0.12)	0.94 (0.13)	0.88 (0.20)	232.20 (65.34)	209.13 (53.34)	179.01 (42.65)	0.12 (0.07)	0.14 (0.12)	0.12 (0.09)
no-hand→pen	0.89 (0.18)	0.80 (0.23)	0.84 (0.22)	203.41 (65.43)	204.13 (77.51)	173.94 (51.98)	0.12 (0.09)	0.16 (0.14)	0.12 (0.11)

Note. Completion times are indicated in seconds.

3.2.3.2 Transfer Effects on Route Planning Task Performance

To test the hypothesis that the decline in performance following the removal of cognitive offloading would be greater than the improvement following its introduction (H1), we conducted a series of 4 (pen→no-pen *vs*. pen→no-hand *vs*. no-pen→pen *vs*. no-hand→pen groups) x 2 (training-consistent *vs.* training-inconsistent test blocks) mixed ANOVAs on each of the performance measures from the route planning task. In the presence of statistically significant group × test block interactions, the alpha level for post-hoc comparisons was adjusted from .05 to .0125, thereby accounting for the four within-group comparisons across test blocks, which represented the key comparisons of interest for our hypothesis. Tukey's corrections were used for post-hoc comparisons where we did not have specific hypotheses.

Proportion of maps without errors

Considering the proportion of maps completed without errors, a main effect of test block, $F(1,133) = 5.05$, $p = .026$, $\eta^2_p = 0.04$, was observed. Also the main effect of group was found to be significant, $F(1,133) = 2.90$, $p = .037$, $\eta^2_p = 0.06$, while the interaction block by group was not, $F(3,133) = 2.43$, $p = .068$, $\eta^2_p = 0.05$. Overall, participants completed a higher proportion of maps without errors in the training-consistent test block, $M = 0.91$, $SD = 0.16$, than in the training-inconsistent test block, $M = 0.87$, $SD = 0.20$. Participants in the no-hand→pen group completed a lower proportion of maps without errors across the two test blocks ($M = 0.82$, $SD = 0.23$) compared to participants in the other groups (pen→no-pen group, $M = 0.91$, $SD = 0.17$; pen→no-hand group, $M = 0.91$, $SD = 0.16$; no-pen→pen group, $M = 0.91$, $SD = 0.17$), but Tukey's post-hoc comparisons revealed that these differences did not reach statistical significance (no-hand→pen group *vs*. pen→no-pen group, $t(133) = -2.39$, $p = .084$; no-hand→pen group *vs*. pen→no-hand group, $t(133) = -2.40$, $p = .083$; no-hand→pen group *vs*. no-pen→pen group, $t(133) = -2.47$, $p = .069$). Even in this case, the results did not change when controlling for the presentation order of the test blocks.[12]

Proportional Deviation above the Shortest Route

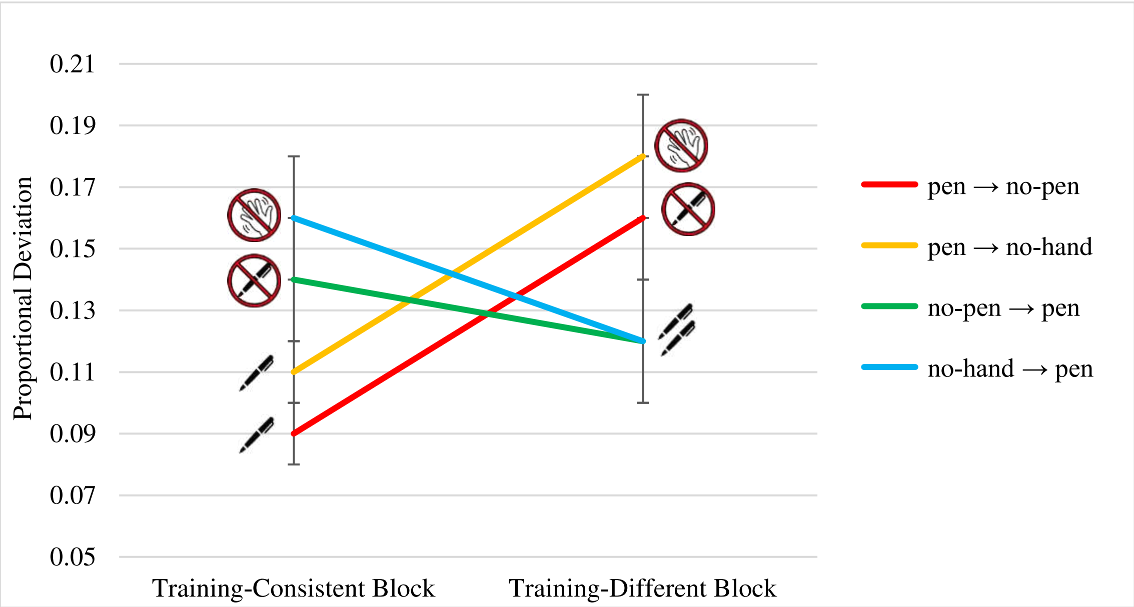
Regarding the proportional deviation above the shortest route, a significant group by test block interaction was found, $F(3,132) = 8.42$, $p = <.001$, $\eta^2_p = 0.16$. However, neither the main effect of group nor the main effect of block reached statistical significance, Fs < 1.88, ps $> .173$. Post-hoc comparisons revealed that both the pen→no-pen group and the pen→no-hand group planned significantly longer routes in the training-inconsistent block than in the training-consistent block (pen→no-pen group: training-consistent block, $M = 0.09$, $SD = 0.08$, training-inconsistent block, $M = 0.16$, $SD = 0.13$, $t(132) = 2.99$, $p = .003$; pen→no-hand group: training-consistent block, $M = 0.11$, $SD = 0.09$, training-inconsistent block, $M = 0.18$, $SD = 0.14$, $t(132) = 3.43$, $p < .001$). In contrast, neither

[12] Main effect of test block, $F(1,129) = 4.87$, $p = .029$, $\eta^2_p = 0.04$. Main effect of group, $F(3,129) = 2.95$, $p = .035$, $\eta^2_p = 0.06$. Tuckey's post-hoc comparisons between groups: no-hand→pen group *vs*. pen→no-pen group, $t(129) = -2.44$, $p = .075$; no-hand→pen group *vs*. pen→no-hand group, $t(129) = -2.40$, $p = .082$; no-hand→pen group *vs*. no-pen→pen group, $t(129) = -2.48$, $p = .068$. All other main and interactive effects were not statistically significant (Fs<2.37, ps>.074).

the no-pen→pen group nor the no-hand→pen group planned significantly shorter routes in the training-inconsistent block compared to the training-consistent block (no-pen→pen group: training-consistent block, $M = 0.14$, $SD = 0.12$, training-inconsistent block, $M = 0.12$, $SD = 0.09$, $t(132) = -1.22$, $p = .225$; no-hand→pen group: training-consistent block, $M = 0.16$, $SD = 0.14$, training-inconsistent block, $M = 0.12$, $SD = 0.11$, $t(132) = -2.31$, $p = .022$) (see Figure 8). Therefore, participants planned significantly longer routes in the test block that precluded pen use for offloading (or both pen and hands) compared to the test block that allowed pen use, depending on their training condition. Specifically, only participants who were previously trained with pen use allowed were significantly affected by the restriction of pen use during the subsequent testing phase. Results did not significantly change after controlling for the presentation order of test blocks.[13]

Figure 8

Transfer Effects on Proportional Deviation Above the Shortest Route



Note. Bars represent standard errors.

[13] Group × test block interaction, $F(3,128) = 8.13$, $p < .001$, $\eta^2p = 0.16$. Post-hoc comparisons between training-inconsistent block vs. training-consistent block: pen→no-pen group, $t(128) = 2.89$, $p = .004$; pen→no-hand group, $t(128) = 3.41$, $p < .001$; no-pen→pen group, $t(128) = -1.21$, $p = .230$; no-hand→pen group: $t(128) = -2.27$, $p = .025$). All other main and interactive effects were not statistically significant ($Fs<1.78$, $ps>.183$).

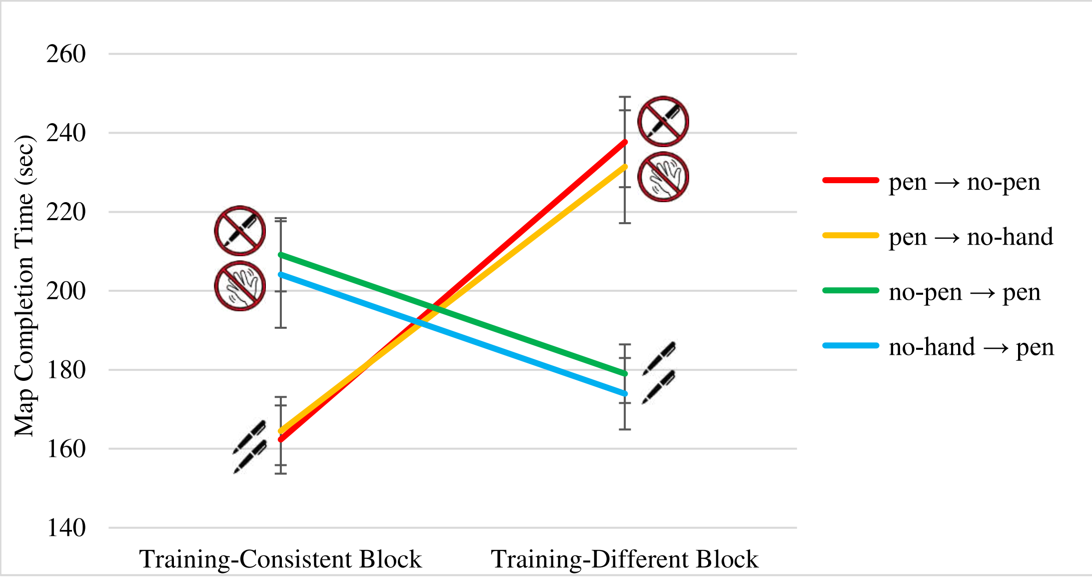
Map Completion Time

The mean time required to complete a map was significantly affected by the test block, $F(1,132) = 16.80$, $p < .001$, $\eta^2_p = 0.11$, while the main effect of group was not statistically significant, $F(3,132) = 0.28$, $p = .843$, $\eta^2_p = 0.01$. The group by test block interaction was also statistically significant, $F(3,132) = 34.66$, $p < .001$, $\eta^2_p = 0.45$. Post-hoc comparisons revealed that participants in both the pen→no-pen group and the pen→no-hand group required more time to complete maps in the training-inconsistent block compared to the training-consistent block (pen→no-pen group: training-consistent block, $M = 162.34$ s, $SD = 51.12$ s; training-inconsistent block, $M = 237.68$ s, $SD = 66.70$ s; $t(132) = 7.53$, $p < .001$; pen→no-hand group: training-consistent block, $M = 164.49$ s, $SD = 51.78$ s; training-inconsistent block, $M = 231.42$ s, $SD = 85.73$ s; $t(132) = 6.96$, $p < .001$). Additionally, participants in the no-pen→pen group and the no-hand→pen group spent significantly less time completing maps in the training-inconsistent block than in the training-consistent block (no-pen→pen group: training-consistent block, $M = 209.13$ s, $SD = 53.34$ s; training-inconsistent block, $M = 179.01$ s, $SD = 42.65$ s; $t(132) = -3.00$, $p = .003$; no-hand→pen group: training-consistent block, $M = 204.13$ s, $SD = 77.51$ s; training-inconsistent block, $M = 173.94$ s, $SD = 51.98$ s; $t(132) = -3.01$, $p = .003$) (see Figure 9). The results did not change after controlling for the presentation order of the test blocks.[14]

[14] Main effect of test block, $F(1,128) = 16.48$, $p < .001$, $\eta^2 p = 0.11$. Group × test block interaction, $F(3,128) = 34.00$, $p < .001$, $\eta^2 p = 0.44$. Post-hoc comparisons between training-inconsistent block vs. training-consistent block: pen→no-pen group, $t(128) = 7.47$, $p < .001$; pen→no-hand group, $t(128) = 6.88$, $p < .001$; no-pen→pen group, $t(128) = -2.97$, $p = .004$; no-hand→pen group: $t(128) = -2.97$, $p = .004$). All other main and interactive effects were not statistically significant (Fs<0.52, ps>.477).

Figure 9

Transfer Effects on Map Completion Time

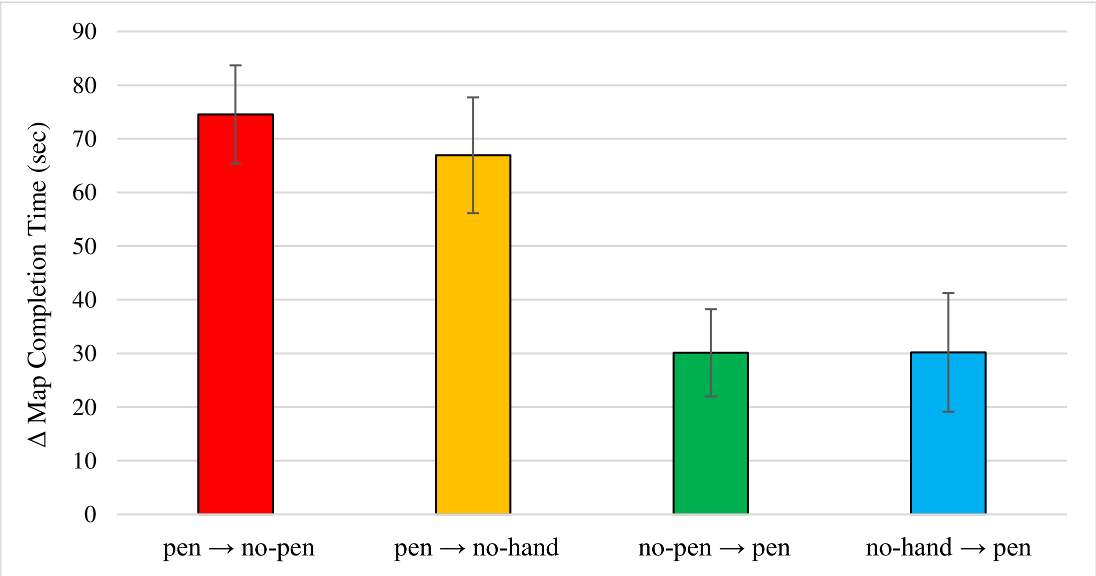


Note. Bars represents standard errors.

Therefore, regardless of their group, participants were significantly slower in planning their routes in the test block that precluded pen use for offloading (or both pen and hands) compared to the test block that allowed pen use. According to our first hypothesis (H1), when comparing the training-inconsistent block to the training-consistent block, the degree of slowdown showed by groups trained with pen use should exceed the degree of speedup exhibited by groups trained without pen use. To test this hypothesis, we first calculated for each group the difference in average map completion times (Δ map completion times) between the test block that precluded pen use (training-inconsistent block for the pen→no-pen and pen→no-hand groups; training-consistent block for the no-pen→pen and no-hand→pen groups) and the test block that allowed pen use (training-consistent block for the pen→no-pen and pen→no-hand groups; training-inconsistent block for the no-pen→pen and no-hand→pen groups). Higher Δ map completion times indicates a greater slowdown after pen removal for groups trained with pen use allowed (pen→no-pen and pen→no-hand groups) and a greater speedup after pen introduction for groups trained with pen use precluded (no-pen→pen and no-hand→pen groups). We then conducted a one-way ANOVA on Δ map completion time, revealing significant between-group differences, $F(3,132) = 5.68$, $p = .001$. Tukey's post-hoc comparisons showed that participants in the pen→no-pen group ($M = 74.54$ s, $SD = 53.35$ s) had significantly higher Δ map completion times than those in both the no-pen→pen group ($M = 30.12$ s, $SD = 46.64$ s), $t(132) = 3.15$, $p = .011$, and the no-hand→pen group ($M = 30.19$ s, $SD = 63.52$ s), $t(132) = 3.15$, $p = .011$. Similarly, participants in the pen→no-hand group ($M = 66.93$ s, $SD = 64.73$ s) had significantly higher Δ map completion times than those in the no-pen→pen group, $t(132) = 3.15$, $p = .011$, and the no-hand→pen group, $t(132) = 3.15$, $p = .011$ (see Figure 10). The remaining between-group comparisons were not statistically significant (ts < 0.78, ps > .865).

Figure 10

Extent of Slowdown/Speedup After the Removal/Introduction of Pen Use as a Function of the Group



Note. Bars represents standard errors.

Overall, these findings align with the asymmetric transfer effects predicted by H1. During the test phase, participants trained with the pen exhibited a greater decline in route planning performance after its removal than the improvement observed in participants trained without the pen after its introduction. Specifically, participants trained with the pen planned significantly less accurate (i.e., longer) routes once the use of pen was removed, while those trained without the pen improved their planning accuracy once given the pen, though this improvement was not significant. Additionally, the slowdown in route planning after pen removal for participants trained with the pen was greater than the speedup observed in participants trained without the pen. However, no asymmetric transfer effects were found in the proportion of maps completed without errors, which remained consistently high across all groups and test blocks.

Our hypothesis regarding these effects was grounded in the assumption that participants trained with the pen would rely on a primarily visual strategy for planning and struggle to adjust to a strategy more based on working memory once the pen was removed. Conversely, participants trained without the pen were expected to shift from a mainly working memory-based strategy to a mainly visual strategy with minimal difficulty once the pen was introduced.[15]

It also follows from this hypothesis that participants trained with the pen should show less performance decline after pen removal if they had higher working memory capacity, as this would facilitate their shift to a strategy more based on working memory. To test this additional hypothesis, we conducted an explorative moderation analysis using General Linear Models (GLM). We focused on map completion time, specifically the difference in average map completion time (Δ map completion times) between test blocks allowing and precluding pen use, as this measure showed the most pronounced asymmetric transfer effects. The GLM included group, spatial working memory score (Symmetry Span), and their interaction as predictors of Δ map completion times. Group was a significant predictor, $F(3,127) = 5.66$, $p = .001$, $\eta^2_p = .12$, consistent with the one-way ANOVA results (see above). The spatial working memory score was not a significant predictor, $F(1,127) = 1.30$, $p = .256$, $\eta^2_p = .01$. Notably, the interaction between group and spatial working memory was statistically significant, $F(3, 127) = 3.65$, $p = .014$, $\eta^2_p = .08$.

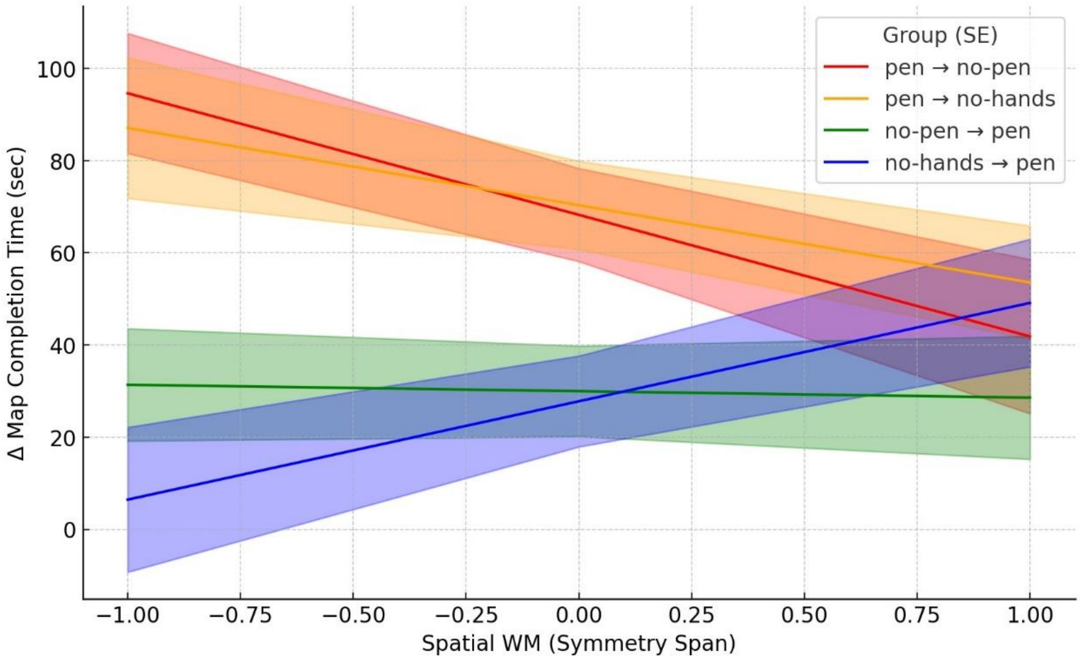
[15] Some exploratory correlation analyses provided further support for these assumptions. Specifically, when pen use was allowed, better visual search abilities were positively correlated with the use of offloading strategies in the training phase, $r(69) = .32$, $p = .006$. Additionally, a greater use of offloading strategies was associated with smaller deviations above the shortest route in the test phase, $r(134) = -.23$, $p = .004$. Conversely, when pen use was precluded, higher spatial working memory was related to smaller deviations above the shortest route in the training phase, $r(64) = -.30$, $p = .004$. This pattern of correlation analyses is generally consistent with those observed in Study 1. For a full report of these correlation analyses, see Appendix B, section 2.

Simple effects analyses showed that groups differed in Δ map completion time when spatial working memory scores were 1 *SD* below the mean, $F(3,127) = 8.98$, $p < .001$, $\eta^2_p = .18$, and at the mean, $F(3,127) = 5.66$, $p = .001$, $\eta^2_p = .12$, but not 1 *SD* above the mean, $F(3, 127) = 0.70$, $p = .56$, $\eta^2 p = .02$. Specifically, when spatial working memory scores were 1 *SD* below the mean, participants in the pen→no-pen group had significantly greater Δ map completion times than those in the no-pen→pen group, $t(127) = 3.54$, $p < .001$, and the no-hand→pen group, $t(127) = 4.32$, $p < .001$, but did not differ significantly from those in the pen→no-hand group, $t(127) = 0.38$, $p = .708$. At the spatial working memory mean, participants in the pen→no-pen group showed higher Δ map completion times than those in the no-pen→pen group, $t(127) = 2.72$, $p = .007$, and the no-hand→pen group, $t(127) = 2.87$, $p = .005$, but not from those in the pen→no-hand group, $t(127) = -0.15$, $p = .881$. When spatial working memory scores were 1 *SD* above the mean, no significant between-group differences in Δ map completion times were observed ($|t|$s < 0.63, *p*s $> .535$) (see Figure 11).

In the transition to the training-different test block, participants trained with the pen experienced a slowdown greater than the speedup exhibited by participants trained without the pen, except when they had higher spatial working memory, consistent with our interpretation of the asymmetric transfer effects.

Figure 11

Extent of Slowdown/Speedup in Map Completion Time After the Removal/Introduction of Pen Use as a Function of Spatial Working Memory.



*Note***.** The values on the X-axis represent Symmetry Span scores, standardized as Z-scores.

3.2.3.3 Transfer Effects on Route Planning Task Perception

Our second hypothesis was that the removal of pen use would lead to an increase in perceived task difficulty and effort, while the introduction of pen use would result in a decrease in these perceptions (H2). To test this hypothesis, a 4 (pen→no-pen *vs*. pen→no-hand *vs*. no-pen→pen *vs*. no-hand→pen groups) x 2 (training-consistent *vs*. training-inconsistent test blocks) mixed ANOVA was performed on the perceived difficulty of the task, and one on perceived effort. When statistically significant group × test block interactions were observed, the alpha level for post-hoc comparisons was adjusted from .05 to .0125 to account for the four key within-group comparisons across test blocks. For post-hoc comparisons without specific hypotheses, Tukey's corrections were applied.

Task perceived difficulty

A significant interaction between group and test block was observed for the perceived difficulty of the route planning task, $F(3, 133) = 58.18$, $p < .001$, $\eta^2_p = 0.57$. The main effect of group was also statistically significant, $F(3, 133) = 3.02$, $p = .032$, $\eta^2_p = 0.06$, while the main effect of test block was not, $F(1, 133) = 2.45$, $p = .120$, $\eta^2_p = 0.02$. Post-hoc comparisons revealed that participants in both the pen→no-pen group and the pen→no-hand group rated the task as more difficult in the training-inconsistent block than in the training-consistent block (pen→no-pen group: training-inconsistent block, $M = 5.57$, $SD = 1.40$; training-consistent block, $M = 3.74$, $SD = 1.24$; $t(133) = 8.01$, $p < .001$; pen→no-hand group: training-inconsistent block, $M = 5.61$, $SD = 0.99$; training-consistent block, $M = 4.11$, $SD = 1.17$; $t(133) = 6.67$, $p < .001$). Conversely, participants in the no-pen→pen group and the no-hand→pen group perceived the task as less difficult in the training-inconsistent block compared to the training-consistent block (no-pen→pen group: training-different block, $M = 4.06$, $SD = 1.14$; training-consistent block, $M = 4.88$, $SD = 1.24$; $t(133) = -3.48$, $p < .001$; no-hand→pen group: training-inconsistent block, $M = 3.30$, $SD = 1.19$; training-consistent block, $M = 5.09$, $SD = 0.88$; $t(133) = -7.61$, $p < .001$).[16]

Task perceived effort

A similar pattern of results was found for task perceived effort. Specifically, the interaction between group and test block was statistically significant, $F(3, 133) = 35.92$, $p < .001$, $\eta^2_p = 0.45$. The main effect of group was also statistically significant, $F(3, 133) = 3.89$, $p = .011$, $\eta^2_p = 0.08$, whereas the main effect of test block was not, $F(1, 133) = 0.05$, $p = .819$, $\eta^2_p = 0.00$. Post-hoc comparisons for the significant group × test block interaction revealed that participants in the pen→no-pen group and the pen→no-hand group reported significantly more effort to complete the training-inconsistent block compared to the training-consistent block (pen→no-pen group: training-inconsistent block, $M = 5.36$,

[16] Group × test block interaction, $F(3,129) = 56.59$, $p < .001$, $\eta^2 p = 0.57$. Main effect of group, $F(1,129) = 3.02$, $p = .032$, $\eta^2_p = 0.07$. Post-hoc comparisons between training-inconsistent block vs. training-consistent block: pen→no-pen group, $t(129) = 7.88$, $p < .001$; pen→no-hand group, $t(129) = 6.59$, $p < .001$; no-pen→pen group, $t(129) = -3.41$, $p < .001$; no-hand→pen group: $t(129) = -7.52$, $p < .001$). All other effects were not statistically significant (Fs < 2.62, ps > .108).

SD = 1.38; training-consistent block, *M* = 4.06, *SD* = 1.43; *t*(133) = 5.33, $p < .001$; pen→no-hand group: training-inconsistent block, *M* = 5.78, *SD* = 1.02; training-consistent block, *M* = 4.39, *SD* = 1.36; *t*(133) = 5.36, $p < .001$). Conversely, participants in the no-pen→pen group and the no-hand→pen group reported putting in significantly less effort in the training-inconsistent block compared to the training-consistent block (no-pen→pen group: training-different block, *M* = 4.00, *SD* = 1.25; training-consistent block, *M* = 5.03, *SD* = 1.29; *t*(133) = -3.81, $p < .001$; no-hand→pen group: training-inconsistent block, *M* = 3.48, *SD* = 1.25; training-consistent block, *M* = 5.12, *SD* = 1.08; *t*(133) = -6.05, $p < .001$). Even in this case, the results did not change when controlling for test block presentation order.[17]

Consistent with the predictions of H2, removing the possibility to offload cognition via pen use during the test phase increased perceived task difficulty and effort among participants who were trained with this option, while introducing the pen led to a decrease in both perceived task difficulty and effort among those who were trained without pen use. Since task perception measures often do not closely align with actual performance measures (e.g., Dunn et al., 2019), we did not explicitly anticipate to observe asymmetric transfer effects in perceived task difficulty and effort. However, we conducted further exploratory analyses to investigate whether the asymmetric effects observed in task performance measures could also be found on task perception measures. For both perceived task difficulty and effort, we followed the Δ score procedure used for map completion times (see section 3.2.3.2). Specifically, we calculated Δ scores for perceived difficulty and perceived effort by subtracting participants' ratings for the test block allowing pen use (the training-consistent block for the pen→no-pen and pen→no-hand groups, and the training-inconsistent block for the no-pen→pen and no-hand→pen groups) from their ratings for the test block precluding pen use (the training-inconsistent block for the pen→no-pen and pen→no-hand groups, and the training-consistent block for

[17] Group × test block interaction, $F(3,129) = 35.10$, $p < .001$, $\eta^2p = 0.45$. Main effect of group, $F(1,129) = 3.85$, $p = .011$, $\eta^2_p = 0.08$. Post-hoc comparisons between training-inconsistent block vs. training-consistent block: pen→no-pen group, $t(129) = 5.26$, $p < .001$; pen→no-hand group, $t(129) = 5.31$, $p < .001$; no-pen→pen group, $t(129) = -3.76$, $p < .001$; no-hand→pen group: $t(129) = -5.97$, $p < .001$). No additional effects were found to be statistically significant. (*F*s < 2.09, *p*s > .151).

the no-pen→pen and no-hand→pen groups). We then conducted two one-way ANOVAs: one for the perceived difficulty Δ scores, and one for the perceived effort Δ scores.

The ANOVA on the perceived difficulty Δ scores revealed a significant main effect of group, $F(3,133) = 3.99, p = .009$. Tukey's post-hoc comparisons indicated that participants in the no-pen→pen group ($M = 0.82, SD = 1.10$) exhibited significantly lower Δ scores than those in both the pen→no-pen group ($M = 1.83, SD = 1.44$), $t(133) = -3.08, p = .013$, and the no-hand→pen group ($M = 1.79, SD = 1.19$), $t(133) = -2.92, p = .021$. No other post-hoc comparisons were statistically significant (*t*s < 2.11, *p*s > .159). The ANOVA on the perceived effort Δ scores did not indicate significant between-group differences, $F(2,133) = 0.85, p = .467$.

Therefore, participants' perceptions of the route planning task in terms of difficulty and effort did not fully align with their actual task performance. The increase in task perceived difficulty experienced by the pen→no-pen group when the pen was removed was greater than the reduction reported by the no-pen→pen group when the pen was introduced. However, the increase in perceived difficulty for the pen→no-hand group was not significantly greater than the reduction experienced by the no-pen→pen group. Interestingly, the results also showed that the decrease in perceived difficulty when the pen was introduced for the first time during the test stage was greater for participants who were trained with hand restriction (no-hand→pen group) compared to those who were trained with pen restriction (no-pen→pen group), suggesting that blocking hands may influence more perceived task difficulty. Regarding perceived task effort, all participants reported exerting less effort during the test block allowing pen use compared to the block precluding it. However, these differences were not more pronounced in participants who were trained with the pen than in those who were trained without it.

3.2.3.4 Use of Offloading Strategies

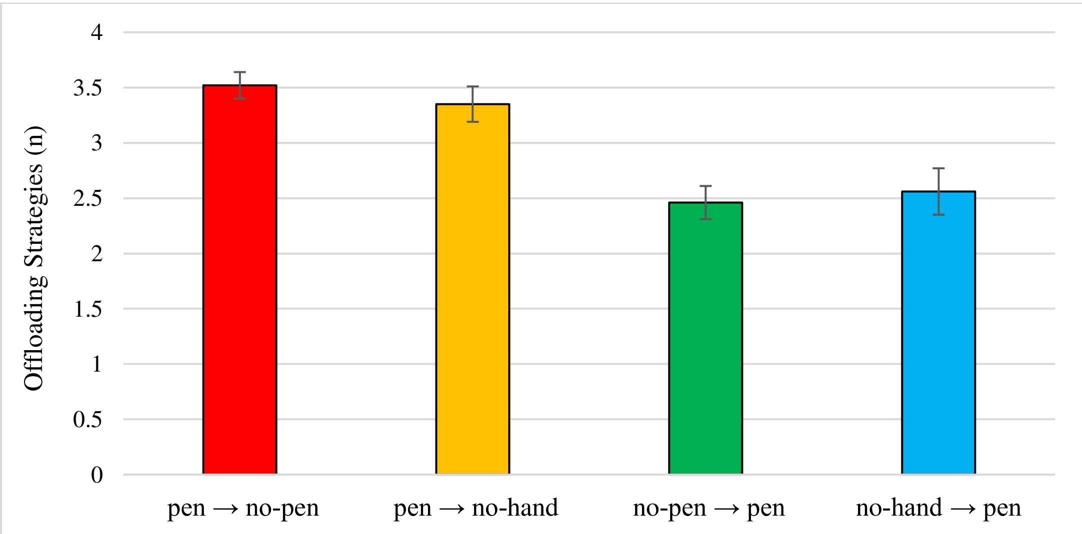
To test the hypothesis that participants trained with pen use, compared with those trained without the pen, would employ more offloading strategies in the test block allowing pen use (H3), we conducted a one-way ANOVA on the average number of strategies devised using the pen in this block.

Groups significantly differ in the average number of offloading strategies used, $F(3, 133) = 11.01$, $p < .001$, $\eta^2_p = 0.20$. Tukey's post-hoc comparisons indicated that participants in the pen→no-pen group ($M = 3.52$, $SD = 0.73$) used more strategies than those in the no-pen→pen group ($M = 2.46$, $SD = 0.87$), $t(133) = 4.58$, $p < .001$, and the no-hand→pen group ($M = 2.56$, $SD = 1.22$), $t(133) = 4.19$, $p < .001$. Similarly, participants in the pen→no-hand group ($M = 3.35$, $SD = 0.94$) used more strategies than those in the no-pen→pen group, $t(133) = 3.86$, $p < .001$, and the no-hand→pen group, $t(133) = 3.47$, $p = .004$. The other post-hoc comparisons were not statistically significant (*t*s < 0.77, *p*s > .871). The results were not different when controlling for test block presentation order.[18] Figure 12 provides a graphical representation of the number of offloading strategies used as a function of the group.

Four additional one-way ANOVAs, each focused on one of the four categories of offloading strategies, were conducted to determine whether specific strategies were more frequently used by participants trained with the pen compared to those trained without it.

Figure 12

Use of Offloading Strategies in the Test Block Allowing Pen Use as a Function of the Group



Note. Bars represent standard errors.

[18] Main effect of group, $F(3,129) = 10.54$, $p < .001$, $\eta^2 p = 0.20$. Post-hoc comparisons: pen→no-pen group *vs*. no-pen→pen group, $t(129) = 4.46$, $p < .001$; pen→no-pen group *vs*. no-hand→pen group $t(129) = 4.07$, $p < .001$; pen→no-hand group *vs*. no-pen→pen group , $t(129) = 3.81$, $p = .001$; pen→no-hand group *vs*. no-hand→pen group, $t(129) = 3.42$, $p = .005$). None of the other effects reached statistical significance (*F*s < 0.31, *p*s > .583).

Groups did not differ in the proportion of test maps with starting and arrival places marked out (pen→no-pen group, $M = 0.96$, $SD = 0.18$, *vs*. pen→no-hand group, $M = 0.92$, $SD = 0.28$, *vs*. no-pen→pen group, $M = 0.87$, $SD = 0.33$, *vs*. pen→no-hand group, $M = 0.79$, $SD = 0.42$), $F(3,133) = 1.94$, $p = .126$, $\eta^2p = 0.04$, nor in the proportion of maps with intermediate locations marked out (pen→no-pen group, $M = 0.80$, $SD = 0.39$, *vs*. pen→no-hand group, $M = 0.83$, $SD = 0.38$, *vs*. no-pen→pen group, $M = 0.59$, $SD = 0.48$, *vs*. pen→no-hand group, $M = 0.67$, $SD = 0.45$), $F(3,133) = 2.54$, $p = .059$, $\eta^2p = 0.05$. However, significant between-group differences were observed in the proportion of test maps with ordering rules marked out, $F(3,133) = 8.92$, $p < .001$, $\eta^2p = 0.17$, and in the proportion of maps with plan progression tracked, $F(3,133) = 5.90$, $p < .001$, $\eta^2p = 0.12$. Tukey's post-hoc comparisons indicated that participants in the pen→no-pen group ($M = 0.88$, $SD = 0.32$) marked the ordering rules in a higher proportion of maps than those in the no-pen→pen group ($M = 0.39$, $SD = 0.44$), $t(133) = 4.75$, $p < .001$, and the no-hand→pen group ($M = 0.54$, $SD = 0.46$), $t(133) = 3.36$, $p = .006$. Participants in the pen→no-hand group ($M = 0.74$, $SD = 0.44$) also marked the rules more frequently than those in the no-pen→pen group, $t(133) = 3.44$, $p = .004$, though not more frequently than those in the no-hand→pen group, $t(133) = 2.04$, $p = .179$. Similarly, participants in the pen→no-pen group ($M = 0.89$, $SD = 0.32$) tracked their plan on a higher proportion of test maps than those in the no-pen→pen group ($M = 0.62$, $SD = 0.45$), $t(133) = 2.80$, $p = .029$, and the no-hand→pen group ($M = 0.57$, $SD = 0.45$), $t(133) = 3.33$, $p = .006$. Participants in the pen→no-hand group ($M = 0.86$, $SD = 0.35$) also tracked their plan progression more frequently than those in the no-hand→pen group, $t(133) = 3.09$, $p = .013$, though the difference was only marginally significant compared to those in the no-pen→pen group, $t(133) = 2.57$, $p = .055$. All other comparisons were far from statistical significance, both for ordering rules marked out (ts < 1.38, ps > .517) and for plan progress tracked (ts < 0.53, ps > .954).

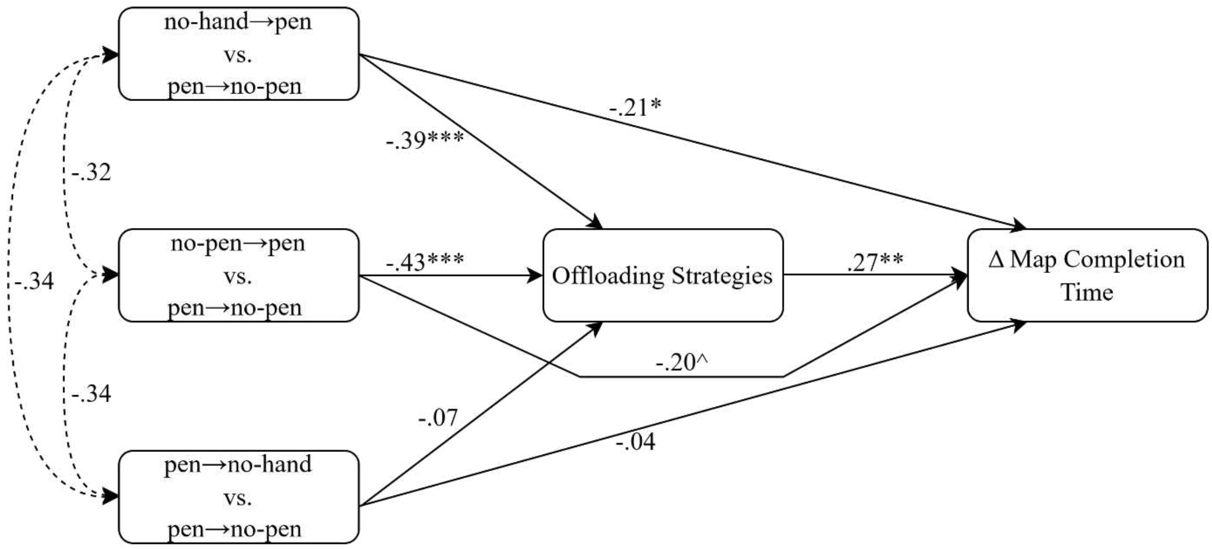
Consistent with hypothesis H3, participants trained with pen use allowed employed a greater number of offloading strategies during the test phase compared to those trained without the pen, indicating a strategy perseveration effect. Specifically, participants trained with pen use allowed relied more frequently on two out of four offloading strategies: marking out ordering rules and tracking plan progression.

Given this strategy perseveration effect, we conducted an additional exploratory mediation analysis to examine whether the number of offloading strategies employed during the testing phase mediated the between-group differences in the observed slowdown or speedup after the pen was precluded or allowed (Δ map completion time). The model illustrated in Figure 13 shows a partial mediation effect, where group membership is related with the utilization of offloading strategies in the test block allowing pen use, which in turn is associated with the Δ map completion time. Specifically, during the test block with pen use allowed, participants in the pen→no-pen group employed significantly more offloading strategies than those in the no-pen→pen group, β = .43, 95% CI [0.60, 1.52], $p < .001$, and the no-hand→pen group, β = .39, 95% CI [0.51, 1.22], $p < .001$, but not more than those in the pen→no-hand group, β = .07, 95% CI [-0.28, 0.62], $p = .449$. This increased reliance on offloading strategies, in turn, resulted in a greater Δ map completion time in the pen→no-pen group compared to both the no-pen→pen group (bootstrapping estimate with 10,000 cycles; indirect effect, β = -.12, 95% CI [-30.901, -4.412], $p = .010$) and the no-hand→pen group (bootstrapping estimate with 10,000 cycles; indirect effect, β = -.11, 95% CI [-29.60, -3.51], $p = .013$). In other words, participants trained with pen use employed more offloading strategies during the test phase allowing the use of the pen compared to those trained without pen use; this greater reliance on offloading strategies, in turn, led to a more pronounced slowdown after pen removal, compared to the speedup seen in participants after pen introduction.

However, the direct effect of group on Δ map completion time was also marginally to fully significant across group comparisons. Specifically, the direct effect was marginally significant when comparing the pen→no-pen and no-pen→pen groups, β = .20, 95% CI [-0.86, 57.20], $p = .057$, and fully significant when comparing the pen→no-pen and no-hand→pen groups, β = .21, 95% CI [0.81, 58.22], $p = .044$. Thus, beyond the use of offloading strategies in the test block allowing pen use, participants who were trained with the pen exhibited a greater Δ map completion time between this test block and the one precluding pen use, compared to participants who were trained without the pen. That is to say, even when controlling for the use of offloading strategies during the test phase, participants who were trained with the pen experienced a greater slowdown when pen use was removed, compared to the speedup showed by participants trained without the pen when they were provided with it during the test phase. This shows the robustness of our results even when controlling for the number of offloading strategies used in test blocks.

Figure 13

Model of the Relationships Between Groups, Number of Offloading Strategies Used, and Differences in Map Completion Time (Δ Map Completion Time) Between the Test Blocks Allowing vs. Precluding Pen Use



^$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 4

Hypotheses Supported and Not Supported in Study 2

STUDY 2 - Spontaneous Offloading Strategies		
H1: Asymmetric Transfer Effects on Route Planning performance	Extent of performance decline after the removal of offloading greater than the extent of improvement after its introduction.	proportion of maps completed without errors: ***not supported***
		proportional deviation above the shortest route: **supported**
		map completion time: **supported**
H2: Transfer Effects on Task Perceptions	Increase in perceived task difficulty and effort after offloading removal; decrease in these perceptions after offloading introduction.	perceived task difficulty: **supported**
		perceived task effort: **supported**
H3: Strategy Perseveration Effect	Greater use of offloading strategies during the test phase by participants in the offloading condition than by participants trained in the no offloading conditions.	**supported**

3.2.4 Discussion

Study 2 aimed to examine whether training in a route planning task with or without the opportunity to offload cognition via pen use would influence subsequent task performance when the opportunity to offload cognition was either removed or introduced. The results revealed asymmetric transfer effects, with participants trained using the pen experiencing a greater decline in performance when offloading was removed compared to the improvement observed in participants trained without the pen when offloading was introduced, consistent with H1. Participants' perceptions of task difficulty and effort also changed in the expected directions when offloading via pen use was removed or introduced (H2), but the transfer effects on task perceptions were not asymmetrical. Additionally, participants trained with the pen employed more offloading strategies in the test block allowing pen use than those trained without the pen, indicating the expected strategy perseveration effect (H3).

The asymmetric transfer effects suggest that the shift from a primarily visual strategy to a strategy more based on working memory is more challenging than the shift in the opposite direction. Participants trained with the pen showed a greater decline in planning accuracy and speed when the pen was removed, compared to the improvement shown by participants trained without the pen when it was introduced. This could be attributed to the fact that participants trained with the pen may have transformed the planning task into a primarily visual one, similar to tasks like the traveling salesman problem, which are efficiently solved using visual strategies that individuals intuitively apply (MacGregor & Chu, 2011; MacGregor & Ormerod, 1996). Supporting this view, participants trained with the pen devised more offloading strategies when they had higher visual search abilities (see note 15, and Appendix B, section 2). Once the pen was removed, however, they could no longer rely on a visual planning strategy, and they needed to shift to an incremental planning strategy (Hayes-Roth & Hayes-Roth, 1979; Fum & Del Missier, 2000; Reitter & Lebiere, 2010) primarily based on working memory. This interpretation was further supported by the observed correlation between spatial working memory and route planning performance when pen use was precluded (see note 15, and Appendix B, section 2).

These findings highlight the differential cognitive consequences of switching between strategies, particularly the difficulty of abandoning an established externalized strategy for an internal one. Notably, the moderation analysis revealed that participants trained with the pen were better able to cope with its removal when they had higher spatial working memory, suggesting that individual differences in cognitive abilities can buffer the negative effects of strategy shifts. This finding also reinforces the idea that working memory supports the route planning task performance when cognitive offloading is not available.

Task perception also shifted alongside performance. Participants trained with the pen perceived the task as more difficult and effortful when offloading was removed, while those trained without the pen found the task easier and less effortful once offloading became available. These changes in perception were consistent with our expectations, as removing the external support likely increased working memory load, making the task subjectively more challenging. In contrast, introducing the pen likely reduced cognitive load for participants previously deprived of offloading, leading to lower perceptions of difficulty and effort.

Interestingly, although no performance differences were observed between participants trained without the pen and those trained without both pen and hand use, the latter group perceived a greater reduction in task difficulty when provided with the pen. This may suggest that the restriction of hand use, while not significantly impacting performance, could have heightened participants' subjective experience of task difficulty. Once this restriction was lifted, participants may have experienced a greater sense of relief or ease, leading to a more pronounced decrease in perceived difficulty compared to those restricted only from pen use.

The results also provide evidence for a strategy perseveration effect. Participants trained with the pen employed more offloading strategies in the test block allowing pen use compared to those trained without the pen. This finding aligns with previous research showing that prior experience with offloading can lead to over-reliance on such strategies (Scarampi & Gilbert, 2020), even when they no longer offer performance benefits (Study 1). The mediation analysis suggested that this perseveration effect may have contributed to some extent to the observed asymmetric transfer effects, by strengthening the effect of training. Specifically, participants trained with the pen used more offloading strategies in the test block allowing pen use, which resulted in a greater slowdown when the pen was removed compared to the speedup experienced by participants introduced to the pen only at this stage. Notably, despite employing more offloading strategies, participants trained with the pen did not outperform those trained without the pen in the test block allowing pen use, as no significant differences were found in route planning accuracy or speed between the groups. Overall, these findings suggest that the more participants trained with the pen relied on offloading strategies, the more the shift to a strategy primarily based on working memory when pen use was removed was challenging. However, the greater slowdown observed in route planning among participants trained with pen use, compared to the speedup experienced by those initially trained without it, persisted even controlling for the number of offloading strategies used during the test phase. This indicate that the initial training contributed to the asymmetric transfer effects observed in route planning performance both directly and indirectly, through its impact on offloading strategy use.

In summary, the present study revealed asymmetric transfer effects on route planning performance, with participants trained using a pen to offloading cognition experiencing a performance decline in route planning when the possibility to offload cognition was removed in the test phase greater than the performance improvement seen in participants who were allowed to offload cognition during the test phase only. This suggests that shifting from an external strategy, mainly visual, to an internal one, mainly based on working memory, as a consequence of the removal of the opportunity to offload cognition is more challenging than the reverse shift, occurring after the opportunity to offload cognition is introduced. Notably, participants trained with offloading strategies experienced a smaller performance decline when offloading was not allowed if they had higher spatial working memory ability, further supporting the idea that planning routes without offloading relies on a working memory-based strategy. Finally, participants initially trained with the pen relied more heavily on offloading strategies during testing, which contributed to a greater performance decline when offloading was removed, likely due to interference from the previous external strategy.

3.3 Study 3

3.3.1 Research Aims and Hypotheses

Study 3 aimed to extend the investigation of transfer effects on route planning by examining the impact of removing or introducing external aids that supported offloading strategies into the maps, thereby eliminating the need for participants to devise such strategies. Participants were randomly assigned to one of five groups and trained on the route planning task under one of three conditions. A first group was trained in the full offloading condition, where all four offloading strategies were supported by external aids embedded into the maps. A second group was trained in the partial offloading condition, in which only two offloading strategies were supported by external aids. The remaining three groups were trained in the no offloading condition, in which no offloading strategies were supported by external aids.

Following the training block, participants completed a test block, during which all but one group performed the route planning task under conditions that differed from their respective training. Specifically, the groups trained in the full offloading and in the partial offloading condition had their external aids for offloading strategies removed, shifting to the no offloading condition. In contrast, two of the groups that were trained under the no offloading condition had external aids implementing offloading strategies introduced into the maps: one group shifted to the full offloading condition, while the other shifted to the partial offloading condition. The final group, serving as a control group, remained in the no offloading condition throughout both the training and test blocks.

We expected to observe asymmetric transfer effects similar to those found in Study 2. When external aids for cognitive offloading were unavailable, participants would rely on an incremental strategy mainly based on working memory to plan routes. Conversely, when external aids were available, participants would adopt an intuitive, more efficient and mainly visual strategy. Shifting between these strategies was expected to produce carryover effects in both directions. However, while the adoption of an intuitive visual strategy after the introduction of external visual aids would be relatively easy, the development of a working memory-based incremental planning strategy after the adoption of the visual strategy was expected to be harder. As a result, similarly to Study 2, we hypothesized that participants deprived of external offloading aids would experience a greater decline in route planning performance compared to the improvement observed in participants shifting to maps with external support for offloading strategies (H1).

Based on findings from Study 2, we also expected that individual differences in cognitive abilities could buffer these asymmetric transfer effects. In the current study, we assessed verbal working memory using the backward digit span (Monaco et al., 2015), rather than spatial working memory as in previous studies, because we wanted to assess a different aspect of working memory associated with an additional potential role of this memory system in our task. Indeed, we anticipated that verbal working memory in our task could be relevant for encoding, maintaining, and reporting the solution path during planning when no offloading strategies were available, due to the need of verbally encoding and maintaining solution paths in chunks of letters to be reported later. Specifically, when no external visual aids are provided, participants should mentally encode and rehearse the route sequence while incrementally planning it, requiring efficient verbal working memory to maintain and report the final route (see also the results on solution writing times in Study 1). Based on our interpretation of the asymmetric transfer effects, we hypothesized that participants initially trained with external offloading strategies would experience a less pronounced slowdown in maintaining and typing the solution after their removal if they had better verbal working memory (H2), mirroring the findings on map completion time in Study 2.

An additional aim of the current study was to investigate whether providing external aids for all four or only two offloading strategies would lead to differences in route planning performance. In Study 2, participants trained with the pen used more offloading strategies than those trained without it during the test block allowing pen use. Specifically, those trained with the pen marked out more frequently ordering rules and kept an external track of their plan, whereas no significant differences were observed between groups in marking out starting and arrival locations or intermediate locations. Importantly, this greater reliance on offloading strategies did not improve route planning performance, suggesting a "more than optimal" use of offloading strategies. This aligned with previous research on cognitive offloading in other domains, which has demonstrated a bias toward the adoption of external strategies even when this did not enhance task performance (e.g., Gilbert et al., 2020; see also Gilbert et al., 2023, and Study 1). Therefore, we hypothesized that participants performing the route planning task under the full offloading condition, where all four strategies were externally supported, would not outperform those in the partial offloading condition, where only the starting and arrival places and intermediate locations were visually marked out. This lack of performance difference was expected in both the training block and the test block (H3).

Finally, similarly to Study 2, we expected that removing external aids for offloading strategies would increase perceived task difficulty and effort, while introducing them would lead to a decrease in these perceptions (H4a). Furthermore, although we did not expect significant performance differences between participants provided with two or all four offloading strategies aided, we hypothesized that differences in task perception could emerge (H4b). Indeed, in Study 2, although no performance differences were observed between participants trained without the pen and those trained without both pen and hand use, the latter group perceived a greater reduction in task difficulty when provided with the pen. This suggested that the additional restriction, despite not impacting performance, heightened the subjective experience of task difficulty. Thus, we deemed it possible that removing or providing all four offloading strategies, compared to only two, might similarly result in a greater sense of relief or burden, leading to more pronounced changes in perceived task difficulty and effort.

3.3.2. Method

3.3.2.1 Participants

An a priori power analysis indicated that a minimum of 160 participants was required to detect medium effect sizes in our experimental design ($f = 0.25$, $\alpha = 0.05$, $1 - \beta = 0.85$). A total of 167 undergraduates participated in the study (85.63% female, age: $M = 20.37$, $SD = 2.96$).[19] The study was approved by the Ethical Committee of the University of Trieste, and participants provided informed consent before completing the experimental session online.

3.3.2.2 Experimental Design

All participants completed an individual experimental session online, consisting of a training block of the route planning task, followed by a test block. Participants were randomly assigned to one of five groups: one group trained with all four offloading strategies aided by the maps (full offloading condition), one group trained with two offloading strategies aided (partial offloading condition), and three groups trained without any aid for offloading strategies (no offloading condition).

After completing the training block, participants moved to the test block. For participants in the groups trained under the full offloading condition and the partial offloading condition, external aids for offloading strategies were removed, requiring participants to complete the test block under the no

[19] Two participants (1.2%) chose not to indicate their gender and seven participants (4.2%) chose not to indicate their age.

offloading condition. In contrast, one group that was trained under the no offloading condition had all four types of external aids for offloading strategies made available during the test block (shift to full offloading), and another group started from the no offloading condition and had two types of external aids available at the test stage (shift to partial offloading). The final group trained under the no offloading condition continued without external any aid for offloading strategies in the test block, serving as the control group.

This setup resulted in a 5 (group: full offloading → no offloading *vs*. partial → no offloading *vs*. no offloading → full offloading *vs*. no offloading → partial offloading *vs*. no offloading → no offloading) × 2 (block: training *vs*. test) mixed experimental design.

3.3.2.3 Procedure

The entire experimental procedure was generated and administered using the Qualtrics XM software. At the start of the experimental session, participants read and provided informed consent before being randomly assigned to one of five groups: 1. Full offloading → no offloading (n = 34); 2. Partial offloading → no offloading (n = 33); 3. No offloading → full offloading (n = 34); 4. No offloading → partial offloading (n = 34); and 5. No offloading → no offloading (n = 32).

Participants read the instructions for the training block of the route planning task, completed a practice map, and then provided performance estimates on two dimensions: route planning accuracy (deviation from the shortest possible route) and route planning speed (average time needed to plan a route). After completing the maps in the training block, participants provided retrospective judgments on the same performance dimensions and rated the difficulty and effort required to complete the training block.

Next, participants received instructions for the test block. Those in the full offloading → no offloading, partial offloading → no offloading, no offloading → full offloading, and no offloading → partial offloading groups were informed that their task conditions would differ from those in the training block, with specific details provided for each group. Participants in the no offloading → no offloading group were informed that the condition of the test block would be identical to that of their training block. After completing the test block, participants performed the backward digit span task (Monaco et al., 2015) and the visual search task (MFTC; Marra et al., 2013). Finally, they answered demographic questions. Figure 14 provides a graphical representation of the procedure followed to administer Route Planning task.

The entire procedure lasted approximately 40 to 50 minutes. A debriefing session was conducted after data collection, during which participants were fully informed of the study's aims and provided with additional details.

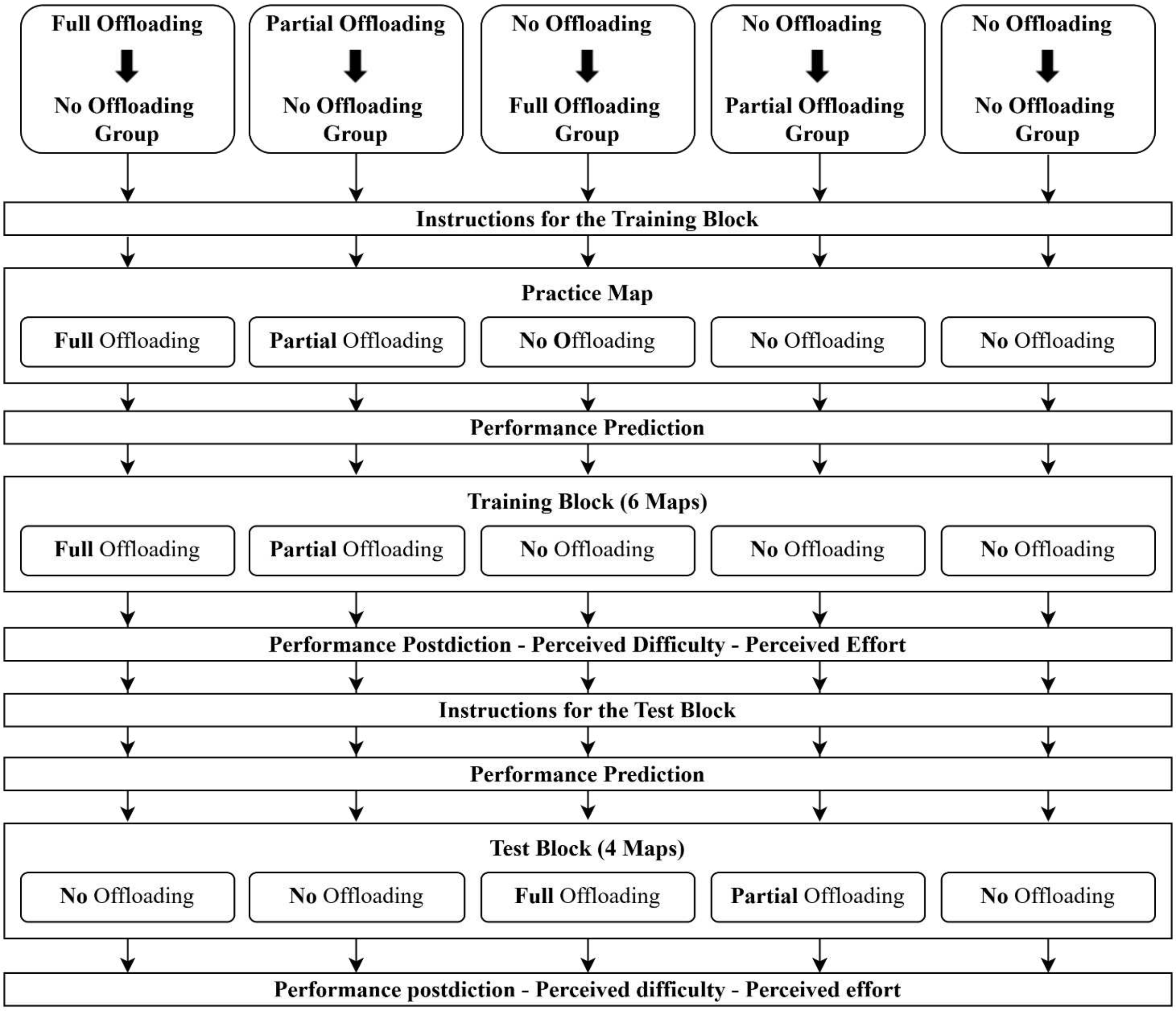
3.3.2.4 Materials

3.3.2.4.1 Route Planning Task

A computerized version of the route planning task was created using the Qualtrics XM software. Similar to Study 2, only high difficulty maps were used, and the training block consisted of 6 maps. However, unlike Study 2, only one test block, consisting of 4 maps, was presented. The decision to include two test blocks in Study 2 was based on findings from Scarampi and Gilbert (2020), who observed that perseveration effects from training (forced offloading *vs*. no offloading) to testing (free-choice) could be eliminated by introducing just one trial using the opposite strategy in an intention offloading task. To account for this, we used two test blocks in Study 2 – one consistent with the training condition and one different – and counterbalanced their order. However, despite this design, strategy perseveration effects from training to the test blocks were not washed out in Study 2. As a result, in the present study, we decided to use only one test block. This change was also aimed at simplifying the procedure and reducing the session length, thereby minimizing participants' fatigue and potentially increasing task compliance in our on-line setting.

Figure 14

Overview of the Route Planning Task Administration Procedure in Study 3



The maps used in this study were a subset of those from Study 2, specifically selected based on their comparability in terms of speed and accuracy when completed under the no-pen condition. The six maps assigned to the training block were overall comparable to the four maps assigned to the test block. Like in Study 2, each map presented a starting point, an ending point, eight intermediate locations to be visited, two ordering rules (each constraining the order of visit of a pair of intermediate locations), and ten distractor locations. To control for potential order effects, the presentation order of maps was randomized both within the training block and the test block.

Participants first read the instructions for the training block and completed a practice map under the same condition as their assigned training condition. Then, they were instructed to plan the shortest route as quickly and accurately as possible. After fully identifying the route for a given map, participants had to click on a "Next" (*Avanti*) button to access the response frame below the map, where they typed the locations (letters) in the exact order they planned to visit them.

The instructions were consistent across all groups, except for the specific offloading strategies visually externalized. Participants in the full offloading → no offloading group were shown an example map that, along with its specific directions on the right (e.g., starting and ending locations, intermediate locations to be visited, and ordering rules to adhere to), presented starting/arrival locations and intermediate locations visually marked, locations constrained by ordering rules connected by arrows, and a textbox for tracking their plan (see Figure 15, panel A). The partial offloading → no offloading group was shown an example map that, along with its specific directions on the right, presented starting/arrival locations and intermediate locations visually marked, but no ordering rule arrows or planning textbox (see Figure 15, panel B). Participants in the no offloading → full offloading, no offloading → partial offloading, and no offloading → no offloading groups were shown an example map with only its specific directions on the right panel and no external visual aids (see Figure 15, panel C).

After reading the instructions, participants completed a practice map. If they made errors (e.g., incorrect starting/arrival points, or violated ordering rules), they were shown their typed route along with feedback specifying the error(s) and they were asked to repeat the practice map. If errors persisted after the second attempt, they were presented with a correct example of a route that met all task criteria, ensuring they understood task requirements before proceeding to the training block.

After completing the six maps in the training block, participants received instructions for the test block. All participants were required to identify the shortest possible routes as quickly and accurately as possible also in the test maps, but in four out of five groups participants were informed that their test condition differed from their training condition. Specifically, participants in the full offloading →

no offloading and partial offloading → no offloading groups were informed that the test maps would no longer incorporate visual aids, with details provided on which aids had been removed based on their training condition. In contrast, participants in the no offloading → full offloading and no offloading → partial offloading groups were informed that the test maps would now include visual aids, with details provided on the specific aids introduced. Finally, participants in the no offloading → no offloading group were informed that the test maps would be presented under the same condition as their training maps. For each map, after typing the entire route, participants had to click the "Next" button to proceed to the following map.

The scoring procedure was identical to that used in the previous studies. First, we measured the proportion of maps completed without errors (i.e., correct starting/arrival locations, inclusion of all intermediate locations, and adherence to ordering rules). For error-free maps, we then assessed planning accuracy by identifying the shortest route and calculating the proportional deviation above this route using the formula: *proportional deviation* = (actual route length - shortest route length) / shortest route length. This metric reflects the quality of the planned route, with higher values indicating greater deviation from the shortest route.

For planning speed, we recorded the time taken to complete each error-free map (the interval between the map's presentation and the participant confirming the typed route). Additionally, the time spent typing the route (the interval between accessing the response frame and confirming the typed route) was measured.

For participants in the full offloading → no offloading and no offloading → full offloading groups, we also recorded the proportion of training and test maps in which participants kept track of their planned route using the online annotations made available.

3.3.2.4.2 Metacognitive Performance Judgments, Perceived Difficulty, and Effort

Performance Prediction and Postdiction. Before both the training and test blocks, participants were asked to estimate how much they would deviate from the shortest possible routes on a 10-point scale (1 = "I will deviate a lot", 10 = "I will deviate very little") and to estimate the average time they would spend planning a route on a 7-point scale (1 = "less than 1 min", 7 = "more than 5 min"). After completing each block, participants provided retrospective judgments on both route planning accuracy and speed using the same response scales.

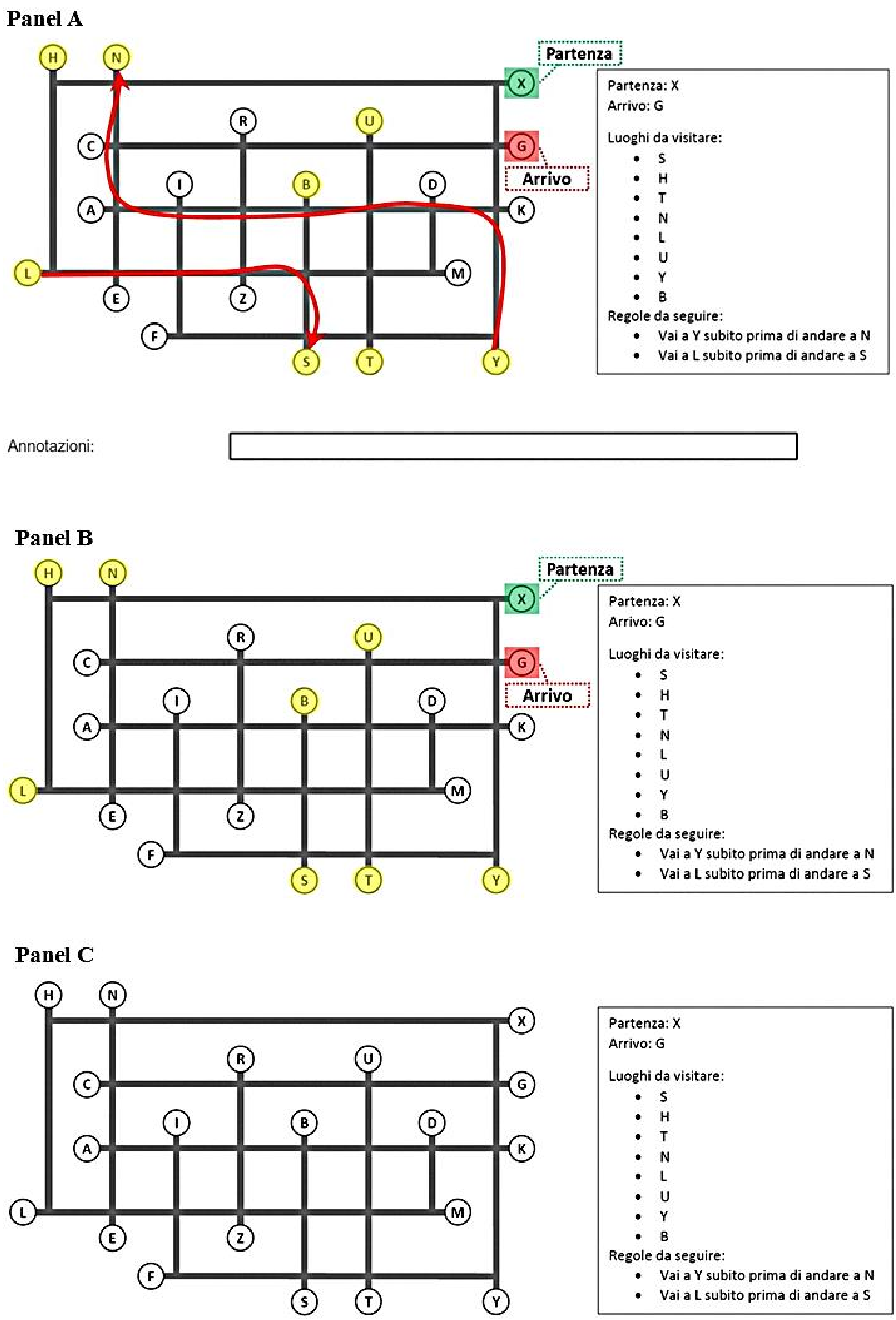
Perceived Difficulty and Effort. Upon completing each block of maps, participants also rated how difficult they found the maps and how much effort they invested in completing them, using a 7-point scale ranging from "very little" to "very much."

3.3.2.4.3 Measures of cognitive abilities

Verbal working memory. Participants completed a computerized version of the Backward Digit Span task adapted from Monaco et al. (2015). In this task, participants were presented with a sequence of digits, each displayed on the screen for one second. The initial span started with two digits. After viewing the sequence, participants were required to type the digits in reverse order. If they responded correctly, the digit span increased by one in the next trial. If they made an error, a new sequence of the same length was presented. Participants were given two attempts to correctly recall a sequence for each span length. If they failed twice at the same span length, their span score corresponded to the maximum number of digits correctly recalled in reverse order, with a maximum possible score of 9.

Figure 15

Example Maps from the Route Planning Task Under Full Offloading (A), Partial Offloading (B), and No Offloading (C) Conditions



Note. In Panel A, the starting place (X – *Partenza*) and arrival place (G – *Arrivo*) are verbally labeled and visually highlighted in green and red, respectively, within the map. Locations to be visited (S, H, T, N, L, U, Y, B – *Luoghi da visitare*) are highlighted in yellow, and ordering rules (Y-N, L-S – *Regole da seguire*) are represented as directed red arrows. A textbox labeled *Annotazioni* below the map provides space to track plan progress. In Panel B, only the starting place (X – *Partenza*), arrival place (G – *Arrivo*), and locations to be visited (S, H, T, N, L, U, Y, B – *Luoghi da visitare*) are visually highlighted. In Panel C, no visual aids for offloading strategies are provided.

Visual search. Participants completed a computerized version of the Multiple Features Target Cancellation task (MFTC; Marra et al., 2013) to assess visual search ability. They viewed a grid of 80 squares, each containing two differently oriented lines, and were tasked with identifying the 13 squares that matched a probe (a square with two oriented lines) displayed above the grid. Timing began when the grid appeared and ended when participants clicked a bottom-right arrow to confirm completion. Accuracy was calculated using the formula: [(correct responses / 13) + ((67 - false alarms) / 67)] / 2.

3.3.3 Results

3.3.3.1 Preliminary analyses

For each performance variable of the route planning task, data were trimmed within groups of participants and blocks of maps. Values that were 3 *SD*s above or below the mean were replaced by the mean ± 3 *SD*s (as in Study 1 and Study 2; see also Del Missier et al., 2021, and Miyake et al., 2000). For verbal working memory and visual search measures, the same replacement procedure was applied across the entire sample. Descriptive statistics for all route planning task measures, stratified by groups and blocks, are presented in Table 5.[20]

[20] Data trimming led to the following percentages of value substitution. For the proportion of maps completed without errors: full offloading → no offloading group, 0.00% for the training block and 0.00% for the test block; partial offloading → no offloading group, 3.03% for the training block and 0.00% for the test block; no offloading → full offloading group, 0.00% for the training block and 2.94% for the test block; no offloading → partial offloading group, 0.00% for the training block and 2.94% for the test block; no offloading →no offloading group, 6.25% for the training block and 3.13% for the test block. For proportional deviation above the shortest route: full offloading → no offloading group, 2.95% for the training block and 0.00% for the test block; partial offloading → no offloading group, 3.03% for both the training and test blocks; no offloading → full offloading group, 5.88% for the training block and 0.00% for the test block; no offloading → partial offloading group, 2.94% for both the training and test blocks; no offloading → no offloading group, 0.00% for the training block and 3.23% for the test block. For map completion time: full offloading → no offloading group, 0.00% for the training block and 3.13% for the test block; partial offloading → no offloading group, 0.00% for both blocks; no offloading → full offloading group, 0.00% for the training block and 3.03% for the test block; no offloading → partial offloading group, 2.94% for the training block and 0.00% for the test block; no offloading → no offloading group, 0.00% for both blocks. For solution-typing time: full offloading → no offloading group, 0.00% for the training block and 3.13% for the test block; partial offloading → no offloading group, 3.03% for both blocks; no offloading → full offloading group, 0.00% for the

There were no significant differences between groups in terms of gender, $\chi^2(4) = 5.55$, $p = .235$, or visual search skills, $F(4,160) = 0.32$, $p = .861$. However, significant between-group differences were found for verbal working memory, $F(4,161) = 3.56$, $p = .008$. Tukey's post-hoc comparisons indicated that, on average, participants in the partial offloading → no offloading group ($M = 6.97$, $SD = 1.96$) had better verbal working memory than those in the full offloading → no offloading group ($M = 5.56$, $SD = 1.40$), $t(161) = 3.56$, $p = .004$. No other between-group differences were statistically significant (ts < 2.68, ps > .062). Given this difference in verbal working memory, subsequent analyses were conducted both with and without controlling for verbal working memory.

Table 5

Descriptive Statistics of Route Planning Performance Measures

	Proportion of Maps Without Errors		**Proportional Deviation Above the Shortest Route**		**Map Completion Time**		**Solution-Typing Time**	
	Training M (SD)	Test M (SD)	Training M (SD)	Test M (SD)	Training M (SD)	Test M (SD)	Training M (SD)	Test M (SD)
full offloading → no offloading	0.89 (0.14)	0.74 (0.30)	0.09 (0.05)	0.18 (0.15)	145.99 (59.86)	266.57 (109.26)	44.06 (26.98)	90.06 (57.47)
partial offloading → no offloading	0.95 (0.11)	0.86 (0.17)	0.10 (0.05)	0.18 (0.14)	140.57 (64.81)	253.11 (102.84)	56.63 (35.64)	85.95 (71.10)
no offloading → full offloading	0.92 (0.12)	0.91 (0.17)	0.11 (0.05)	0.09 (0.06)	204.03 (80.10)	141.66 (53.22)	54.92 (46.89)	34.65 (16.38)
no offloading → partial offloading	0.89 (0.11)	0.93 (0.13)	0.13 (0.07)	0.11 (0.05)	208.78 (67.94)	159.04 (64.36)	54.27 (39.13)	39.34 (20.83)
no offloading → no offloading	0.91 (0.20)	0.84 (0.17)	0.11 (0.07)	0.14 (0.10)	236.32 (85.87)	217.98 (81.71)	64.19 (50.01)	59.23 (40.99)

Note. Map completion times and solution-typing times are indicated in seconds.

training block and 3.03% for the test block; no offloading → partial offloading group, 2.94% for both blocks; no offloading → no offloading group, 3.23% for both blocks. For verbal working memory (Backward Digit Span), there were no substitutions (0.00%). For visual search (MFTC accuracy scores), there were also no substitutions (0.00%). One participant (0.01%), specifically from the no offloading → no offloading group, did not complete the Backward Digit Span task. Two participants (0.01%), one from the partial offloading → no offloading group and one from the no offloading → no offloading group, did not complete the MFTC task. Descriptive statistics for the Backward Digit Span were as follows: $M = 6.09$, $SD = 1.67$. Descriptive statistics for the MFTC accuracy score were as follows: $M = 0.83$, $SD = 0.14$.

3.3.3.2 Transfer Effects on Route Planning Task Performance

We conducted a series of 5 (full offloading → no offloading *vs*. partial offloading → no offloading *vs*. no offloading → full offloading *vs*. no offloading → partial offloading *vs*. no offloading → no offloading groups) × 2 (training *vs*. test blocks) mixed ANOVAs to test the hypothesis that the decline in performance following the removal of externalized offloading strategies would be greater than the improvement following their introduction (H1). For statistically significant group × block interactions, the alpha level for post-hoc comparisons was adjusted from .05 to .01 to account for the five key within-group comparisons.

When both groups trained with external aids for offloading strategies (i.e., full offloading → no offloading and partial offloading → no offloading) exhibited a significant decline in performance after their removal, and both groups trained without aids for offloading strategies (i.e., no offloading → full offloading and no offloading → partial offloading) showed a significant improvement after their introduction, we computed differential scores (Δ) on that measure between the training and test blocks, as was done in Study 2. This allowed us to assess whether the extent of the decline following the removal of external aids for offloading strategies was greater than the improvement following their introduction.

Tuckey's post-hoc comparisons from the above mixed ANOVAs were also used to investigate whether no significant differences in route planning performance would emerge between participants provided with aids for all four offloading strategies and participants provided with aids for two of them only (H3).

Planned comparisons were conducted for the main measures of route planning accuracy (i.e., proportional deviation from the shortest route) and speed (i.e., map completion time), specifically contrasting participants in the full offloading → no offloading and partial offloading → no offloading groups with participants in the no offloading → no offloading group. These comparisons were designed to further explore the impact of prior (*vs*. no prior) experience with aids for offloading on performance when no aids for offloading strategies were available during the test block. The contrast was specified in Linear Mixed Models (LMMs), with group and block as fixed effects and participants' slope included as a random effect. The Satterthwaite method was used to estimate the degrees of freedom.

Proportion of maps completed without errors

For the proportion of maps completed without errors, significant main effects were found for group, $F(4,162) = 2.70$, $p = .033$, $\eta^2_p = .06$, and block, $F(1,162) = 9.63$, $p = .002$, $\eta^2_p = .06$, with a significant group × block interaction, $F(4,162) = 3.64$, $p = .007$, $\eta^2_p = .08$. Post-hoc comparisons revealed that participants in the full offloading → no offloading group completed a significantly higher proportion of maps without errors in the training block ($M = 0.89$, $SD = 0.14$) than in the test block ($M = 0.74$, $SD = 0.30$), $t(162) = 3.89$, $p < .001$. However, participants in the other groups did not exhibit significant differences in the proportion of maps completed without errors between the training and test blocks ($ts < 2.22$, $ps > .029$), with overall performance remaining high across both blocks for all groups (see Table 2). Participants in the full offloading → no offloading and partial offloading → no offloading groups did not significantly differ in the proportion of maps completed without errors in the training block, $t(162) = 1.75$, $p = .767$. Similarly, participants in the no offloading → full offloading group and those in the no offloading → partial offloading group did not show a significant difference in the proportion of maps completed without errors in the test block, $t(162) = 0.27$, $p = 1.000$. After controlling for verbal working memory, the main effect of block was no longer significant, $F(1,160) = 2.00$, $p = .159$, $\eta^2p = .01$, while all other results remained unchanged.[21]

Proportional deviation above the shortest route

Regarding the proportional deviation above the shortest route, the main effect of block, $F(1,158) = 13.61$, $p < .001$, $\eta^2_p = .08$, and the group × block interaction, $F(4,158) = 9.89$, $p < .001$, $\eta^2p = .20$, were both significant, while the main effect of group was not, $F(4,158) = 1.48$, $p = .209$, $\eta^2p = .04$.

[21] Main effect of group, $F(4,160) = 2.44$, $p = .049$, $\eta^2_p = .06$. Group × block interaction, $F(4,160) = 4.08$, $p = .004$, $\eta^2p = .09$. Post-hoc comparisons between the training block and test block for the full offloading → no offloading group, $t(160) = 4.35$, $p < .001$. Post-hoc comparisons between the training block and test block were not significant for all other groups ($ts < 1.71$, $ps > .090$). Tuckey's post-hoc comparison between the full offloading → no offloading and partial offloading → no offloading groups in the training block, $t(160) = 1.04$, $p = .989$. Tuckey's post-hoc comparison between the no offloading → full offloading and no offloading → partial offloading groups in the test block, $t(160) = 0.25$, $p = 1.000$.

Post-hoc comparisons indicated that participants in both the full offloading → no offloading group and the partial offloading → no offloading group planned routes that deviated significantly more from the shortest routes in the test block than in the training block (full offloading → no offloading group: training block, *M* = 0.09, *SD* = 0.05; test block, *M* = 0.18, *SD* = 0.15; $t(158) = 4.78$, $p < .001$; partial offloading → no offloading group: training block, *M* = 0.10, *SD* = 0.05; test block, *M* = 0.18, *SD* = 0.15; $t(158) = 4.74$, $p < .001$). In contrast, neither group provided with aids for offloading strategies in the test block (i.e., no offloading → full offloading group and no offloading → partial offloading group) nor the control group (no offloading → no offloading group) exhibited significant differences in deviation above the shortest routes between the training and test blocks (*t*s < 1.65, *p*s > .102). Furthermore, participants in the full offloading → no offloading group did not significantly differ from participants in the partial offloading → no offloading group in their deviation above the shortest route during the training block, $t(158) = 0.34$, $p = 1.000$, nor did participants in the no offloading → full offloading group differ significantly from those in the no offloading → partial offloading group in the test block, $t(158) = 0.38$, $p = .998$). Results did not significantly change after controlling for verbal working memory.[22]

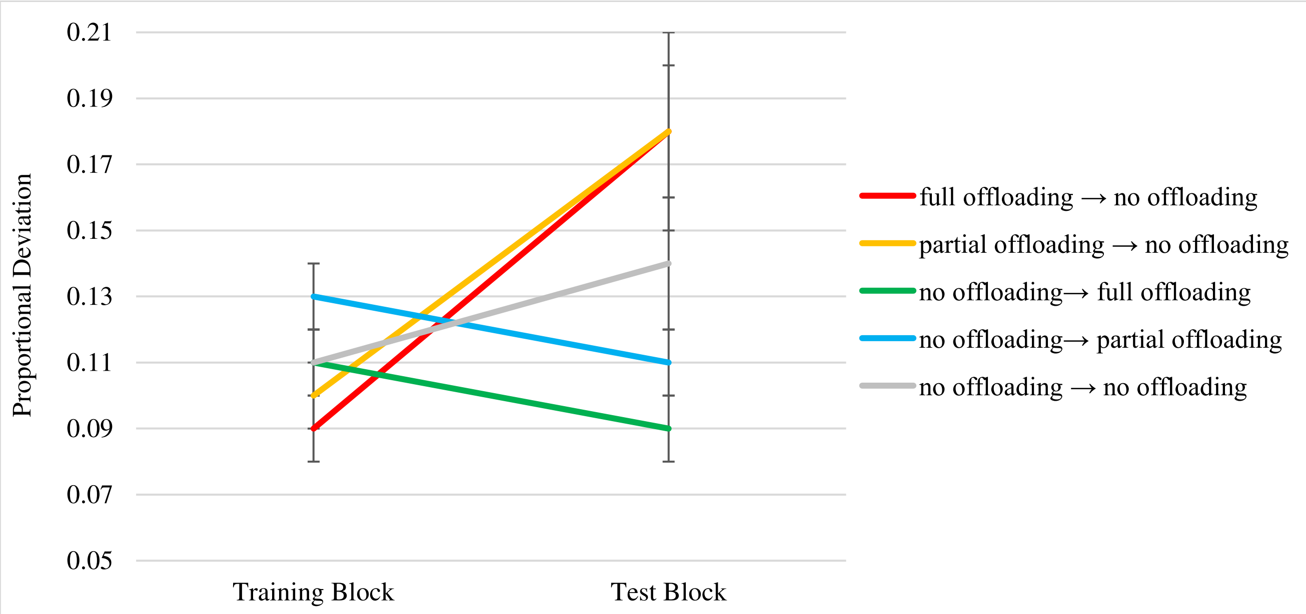
The planned comparison contrasting participants in the full offloading → no offloading and in the partial offloading → no offloading groups with participants in the no offloading → no offloading group on their proportional deviation above the shortest route in the test block was not significant without controlling for verbal working memory, $t(289.41) = 1.86$, $p = .065$. However, after controlling for verbal working memory, the comparison became significant, $t(286.93) = 2.14$, $p = .033$, indicating that participants in the full offloading → no offloading and partial offloading → no offloading groups

[22] Main effect of group, $F(4,156) = 1.66$, $p = .161$, $\eta^2p = .04$. Main effect of block, $F(1,156) = 13.93$, $p < .001$, $\eta^2p = .08$. Group × block interaction, $F(4,156) = 10.99$, $p < .001$, $\eta^2p = .22$. Post-hoc comparisons between the training block and test block for the full offloading → no offloading group, $t(156) = 4.35$, $p < .001$. Post-hoc comparisons between the training block and test block for the partial offloading → no offloading group, $t(156) = 5.40$, $p < .001$. Post-hoc comparisons between the training block and test block were not significant for the remaining groups (*t*s < 1.65, *p*s > .101). Tuckey's post-hoc comparison between the full offloading → no offloading and partial offloading → no offloading groups in the training block, $t(156) = 0.25$, $p = 1.000$. Tuckey's post-hoc comparison between the no offloading → full offloading and no offloading → partial offloading groups in the test block, $t(156) = 0.72$, $p = .999$.

deviated more from the shortest route in the test block than those in the no offloading → no offloading group. See Figure 16 for a graphical representation of the results regarding the proportional deviation above the shortest route.

Figure 16

Transfer Effects on Proportional Deviation above the Shortest Route.



Note. Bars represent standard errors.

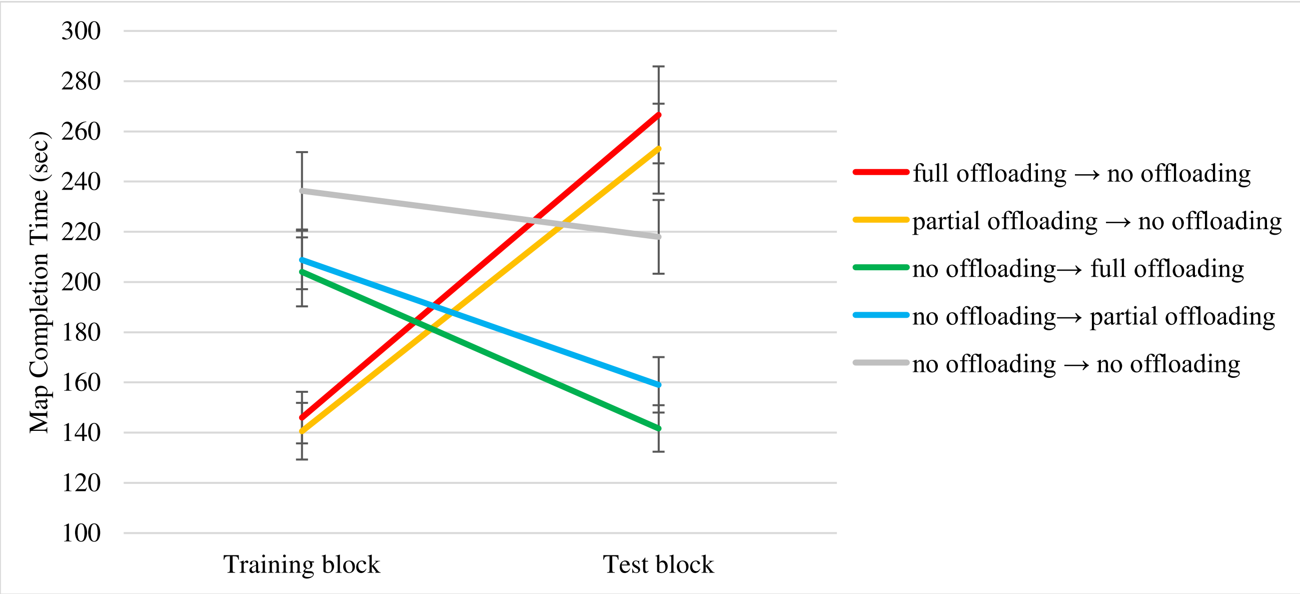
Map completion time

The mean time required to complete a map was significantly affected by both group, $F(4,158) = 3.17$, $p = .013$, $\eta^2p = 0.08$, and block, $F(1,158) = 10.09$, $p = .002$, $\eta^2p = 0.06$. The group × block interaction was also significant, $F(4,158) = 37.53$, $p < .001$, $\eta^2p = 0.49$. Post-hoc comparisons indicated that participants in both the full offloading → no offloading and partial offloading → no offloading groups required significantly more time to complete a map in the test block than in the training block (full offloading → no offloading group: training block, $M = 145.99$ s, $SD = 59.86$ s; test block, $M = 266.57$ s, $SD = 109.26$ s; $t(158) = 8.10$, $p < .001$; partial offloading → no offloading group: training block, $M = 140.57$ s, $SD = 64.81$ s; test block, $M = 235.11$ s, $SD = 102.84$ s; $t(158) = 7.89$, $p < .001$). Conversely, participants in the no offloading → full offloading and no offloading → partial offloading groups spent significantly less time completing maps in the test block than in the training block (no offloading → full offloading group: training block, *M* = 204.03 s, *SD* = 80.10 s; test block, *M* = 141.66 s, *SD* = 53.22 s; *t*(158) = -4.20, *p* < .001; no offloading → partial offloading group: training block, *M* = 208.78 s, *SD* = 67.94 s; test block, *M* = 159.04 s, *SD* = 64.36 s; *t*(158) = -3.54, *p* < .001), while participants in the no offloading → no offloading group did not exhibit significant differences in map completion time between the training block (*M* = 236.32 s, *SD* = 85.87 s) and the test block (*M* = 217.98 s, *SD* = 81.71 s), *t*(158) = 1.25, *p* = .214. In the training block, participants in the full offloading → no offloading group did not complete the task significantly faster than those in the partial offloading → no offloading group, *t*(158) = 0.49, *p* = 1.000. In the test block, participants in the no offloading → full offloading group were not significantly faster than those in the no offloading → partial offloading group, *t*(158) = 0.84, *p* = .998. After controlling for verbal working memory, the main effect of block was no longer significant, *F*(1, 156) = 1.66, *p* = .199, η^2p = 0.01, while the other results remained unchanged.[23]

The planned comparison contrasting participants in the full offloading → no offloading and partial offloading → no offloading groups with participants in the no offloading → no offloading group on map completion time in the test block was statistically significant, *t*(265.61) = 2.40, *p* = .017, and remained significant after controlling for verbal working memory, *t*(262.44) = 2.22, *p* = .027. These results mirror those of the proportional deviation, showing that participants in the full offloading → no offloading and partial offloading → no offloading groups were significantly slower in planning their routes in the test block compared to participants in the no offloading → no offloading group. Figure 17 provides a graphical representation of the results on map completion time.

[23] Main effect of group, *F*(4,156) = 3.42, *p* = .010, η^2p = .08. Group × block interaction, *F*(4,156) = 37.05, *p* < .001, η^2p = .49. Post-hoc comparisons between the training block and test block: full offloading → no offloading group, *t*(156) = 7.87, *p* < .001; partial offloading → no offloading group, *t*(156) = 7.75, *p* < .001; no offloading → full offloading group, *t*(156) = 4.17, *p* < .001; no offloading → partial offloading group, *t*(156) = 3.55, *p* < .001; no offloading → no offloading group, *t*(156) = 1.15, *p* = .251. Tuckey's post-hoc comparison between the full offloading → no offloading and partial offloading → no offloading groups in the training block, *t*(156) = 0.35, *p* = 1.000. Tuckey's post-hoc comparison between the no offloading → full offloading and no offloading → partial offloading groups in the test block, *t*(156) = 0.80, *p* = .998.

Figure 17. Transfer Effects on Map Completion Time.



Note. Bars represents standard errors.

To test whether the slowdown exhibited by participants in the full offloading → no offloading and partial offloading → no offloading groups exceeded the speedup shown by participants in the no offloading → full offloading and no offloading → partial offloading groups, we calculated the difference in map completion time (Δ map completion time) between the training and test blocks for each of these groups.[24] Δ map completion time was computed as the difference between the block with external aids for offloading strategies (training block for the full offloading → no offloading and partial offloading → no offloading groups; test block for the no offloading → full offloading and no offloading → partial offloading groups) and the block without external aids for offloading strategies (training block for the no offloading → full offloading and no offloading → partial offloading groups; test block for the full offloading → no offloading and partial offloading → no offloading groups). A higher Δ map completion time indicates a greater slowdown after the removal of external aids for offloading strategies for the full offloading → no offloading and partial offloading → no offloading groups, and a greater speedup after their introduction for the no offloading → full offloading and

[24] The no offloading → no offloading group was not included in this analysis, as participants in this group did not exhibit significant differences in map completion times between the training and test blocks.

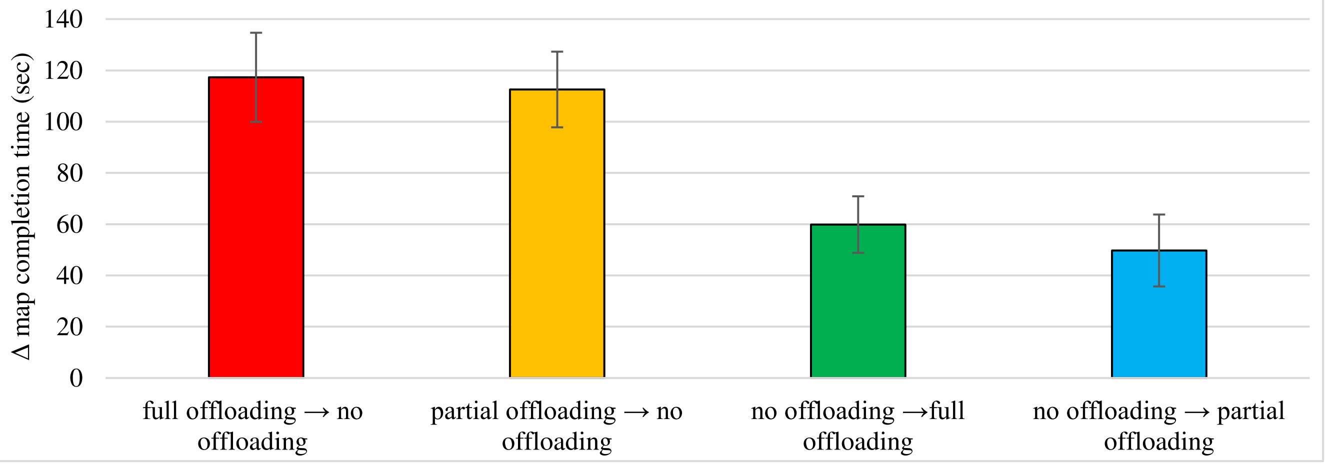
no offloading → partial offloading groups. A one-way ANOVA on Δ map completion time revealed significant between-group differences, $F(3,128) = 5.90$, $p < .001$. Tukey's post-hoc comparisons showed that participants in the full offloading → no offloading group ($M = 117.30$ s, $SD = 98.29$ s) had significantly higher Δ map completion times than participants in both the no offloading → full offloading group ($M = 59.85$ s, $SD = 63.38$ s), $t(128) = 2.79$, $p = .030$, and the no offloading → partial offloading group ($M = 49.74$ s, $SD = 81.78$ s), $t(128) = 3.31$, $p = .007$. Participants in the partial offloading → no offloading group ($M = 112.54$ s, $SD = 84.78$ s) showed marginally higher Δ map completion times compared to participants in the no offloading → full offloading group, $t(128) = 2.58$, $p = .053$, but significantly higher Δ map completion times compared to participants in the no offloading → partial offloading group, $t(128) = 3.10$, $p = .013$. The remaining between-group comparisons were not statistically significant (ts < 0.51, ps > .958), further supporting the idea that removing/introducing aids for either two or four offloading strategies did not differentially increase/decrease map completion time. Notably, after controlling for verbal working memory, the comparison between the partial offloading → no offloading group and the no offloading → full offloading group became fully significant, t(127) = 2.68, p = .041, while all other results remained unchanged. See Figure 18 for a graphical representation of Δ map completion time as a function of the group.

Overall, the results on performance measures are consistent with the predicted asymmetric transfer effects (H1). Participants trained with external aids for offloading strategies (full offloading → no offloading and partial offloading → no offloading groups) exhibited a greater decline in route planning performance when aids for those strategies were removed, compared to the improvement observed in participants who were introduced to external aids for offloading strategies in the test stage (no offloading → full offloading and no offloading → partial offloading groups). Specifically, this was observed in both the proportional deviation above the shortest route and map completion time. Participants in the full offloading → no offloading and partial offloading → no offloading groups deviated significantly more in the test block compared to their performance in the training block. In contrast, participants who were introduced to external aids for offloading strategies in the test stage only did not exhibit a significant improvement in planning accuracy, possibly because their accuracy was already good even without externalized offloading strategies. Participants trained with external aids for offloading strategies showed a significantly greater slowdown in map completion time compared to the speedup observed in participants who were introduced to these aids during the test stage.

Notably, participants in the full offloading → no offloading and partial offloading → no offloading groups were less accurate and slower in the test block compared to those in the no offloading → no offloading group, suggesting that prior experience with (*vs*. without) external aids for offloading strategies negatively affected subsequent performance once these aids were no longer available. However, no significant differences were observed in the proportion of maps completed without errors, with overall performance remaining consistently high across both the training and test blocks for all groups but the full offloading → no offloading group. Participants provided with aids for all four offloading strategies did not outperform those provided with aids for two strategies only, either in the training block or the test block, in line with H3.

Figure 18

Extent of Slowdown/Speedup After the Removal/Introduction of External Aids for Offloading Strategies as a Function on the Group



Note. Bars represents standard errors.

3.3.3.3 Role of Verbal Working Memory on Solution-Typing Time

According to H2, participants initially trained with external visual aids for offloading strategies should experience a less pronounced slowdown in solution-typing time after the removal of these aids if they had better verbal working memory. To test this, we focused on Δ solution-typing time between the training and test blocks. Similarly to the Δ map completion time (see section 3.3.3.2), a higher Δ solution-typing time indicates a greater slowdown for the full offloading → no offloading and partial offloading → no offloading groups after the removal of external aids for offloading strategies, and a greater speedup for the no offloading → full offloading and no offloading → partial offloading groups after their introduction. A General Linear Model (GLM) was used, with group, verbal working memory score (Backward Digit Span), and their interaction as predictors of Δ solution-typing time.[25] Group was a significant predictor, $F(3,124) = 2.82$, $p = .041$, $\eta^2_p = .06$, while verbal working memory score was not, $F(1,124) = 2.84$, $p = .092$, $\eta^2_p = .02$. However, the interaction between group and verbal working memory was statistically significant, $F(3,127) = 2.92$, $p = .037$, $\eta^2_p = .07$. Simple effects analyses revealed that group differences in Δ solution-typing time were significant when verbal working memory scores were 1 *SD* below the mean, $F(3,124) = 5.84$, $p < .001$, $\eta^2_p = .12$, and at the mean, $F(3,124) = 2.81$, $p = .042$, $\eta^2_p = .06$, but not when scores were 1 *SD* above the mean, $F(3,124) = 0.41$, $p = .750$, $\eta^2_p = .01$. Specifically, when verbal working memory scores were 1 *SD* below the mean, participants in the full offloading → no offloading group had significantly greater Δ solution-typing times than those in the no offloading → full offloading group, $t(124) = 3.23$, $p = .002$, and the no offloading → partial offloading group, $t(124) = 3.05$, $p = .003$. No significant difference was found between the full offloading → no offloading and partial offloading → no offloading groups, $t(124) = 0.50$, $p = .961$.

At the mean verbal working memory score, participants in the full offloading → no offloading group did not show significantly higher Δ solution-typing times than those in the

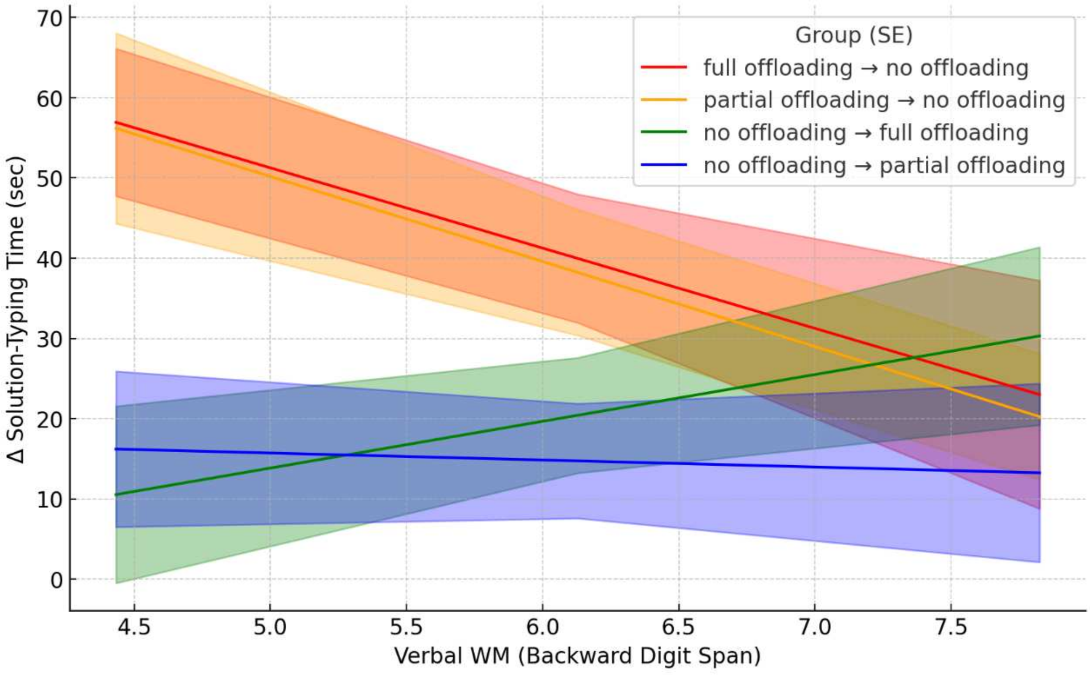
[25] The no offloading → no offloading group was not included in this analysis, as participants in this group did not exhibit significant differences in typing-solution times between the training and test blocks

no offloading → full offloading group, $t(124) = 1.82$, $p = .072$, but they did exhibit significantly higher Δ solution-typing times than those in the no offloading → partial offloading group, $t(124) = 2.35$, $p = .020$. No significant differences were found between the full offloading → no offloading and partial offloading → no offloading groups, $t(124) = 0.16$, $p = .876$.

Importantly, when verbal working memory scores were 1 *SD* above the mean, no significant between-group differences in Δ solution-typing times were observed (*t*s < 0.55, *p*s > .589) (see Figure 19).

Figure 19

Extent of Slowdown/Speedup in Solution-Typing Time After the Removal/Introduction of Offloading Strategies as a Function of Verbal Working Memory



During the transition to the test block, participants trained with aided offloading strategies experienced a greater slowdown in solution-typing time compared to the speedup observed in participants trained without these aids (consistent with H1). This effect was evident for participants with lower and, partially, for those with average verbal working memory, but not for participants with higher verbal working memory, in line with H2.

3.3.3.4 Proportion of Maps with Plan Progression Tracked

The tracking of plan progression was the only offloading strategy that was not fully aided in the full offloading condition. Indeed, although this strategy did not need to be spontaneously devised, it was visually externalized as a textbox below the maps, giving participants the opportunity to take annotations on their route while planning it (e.g., progressively writing the locations to visit in the planned order). Thus, participants were free to choose whether to use this textbox to externally track their route.

Given the strategy perseveration effects found in Studies 1 and 2, as well as in previous research on cognitive offloading (Scarampi & Gilbert, 2020), we explored whether participants in the full offloading → no offloading group took annotations on a greater proportion of maps in the training block compared to participants in the no offloading → full offloading group in the test block. As the Shapiro-Wilk test indicated that the distribution for the proportion of maps with annotations deviated from normality, $W(66) = 0.70$, $p < .001$, we used a Mann-Whitney U test. The analysis revealed that participants in the full offloading → no offloading group took annotations on a significantly greater proportion of maps ($M = 0.61$, $SD = 0.45$, $Median = 1$, $IQR = 0.94$) in the training block compared to participants in the no offloading → full offloading group in the test block ($M = 0.29$, $SD = 0.45$, $Median = 0$, $IQR = 1$), $U = 366.00$, $p = .004$, $r = 0.37$.

Therefore, participants who began the route planning task in the full offloading condition kept a track of their plans more frequently compared to participants who carried out the task in the full offloading condition in the test phase. As the route planning performance of participants in the full offloading → no offloading group during the training block was not superior to the one of participants in the no offloading → full offloading group in the test block (proportional deviation above the shortest route: full offloading → no offloading group, $M = 0.09$, $SD = 0.05$; no offloading → full offloading group, $M = 0.09$, $SD = 0.06$, $t(156) = 0.27$, $p = 1.000$; map completion time: full offloading → no offloading group, $M = 159.99$ s, $SD = 59.86$ s; no offloading → full offloading group, $M = 141.66$ s, $SD = 52.22$ s), this finding provides additional support for H3.[26]

3.3.3.5 Transfer Effects on Route Planning Task Perception

To investigate whether the removal of external aids for offloading strategies led to an increase in perceived task difficulty and effort, and their introduction to a decrease in these perceptions (H4a), we performed two 5 (full offloading → no offloading *vs*. partial offloading → no offloading *vs*. no offloading → full offloading *vs*. no offloading → partial offloading *vs*. no offloading → no offloading groups) × 2 (training *vs*. test blocks) mixed ANOVAs, one for each self-report measure. For statistically significant group × block interactions, the alpha level for post-hoc comparisons was adjusted from .05 to .01 to account for the five key within-group comparisons. To explore whether removing or providing external aids for all four offloading strategies, compared to only two, resulted in more pronounced changes in perceived task difficulty and effort (H4b), the relevant Tukey's post-hoc comparisons from the same mixed ANOVAs were used.

Similar to the rationale followed for route planning performance measures, if both groups trained with external aids for offloading strategies (i.e., full offloading → no offloading and partial offloading → no offloading) experienced a significant increase in perceived difficulty or effort after their removal, and both groups trained without external aids for offloading strategies (i.e., no offloading → full offloading and no offloading → partial offloading) experienced a significant relief after their introduction, we computed differential scores (Δ) on the task perception measures between the training and test blocks. This was done to explore whether the groups exhibited different degrees of change in task perception between the training and the test blocks.

Additionally, we conducted planned comparisons, specifically contrasting participants in the full offloading → no offloading and partial offloading → no offloading groups with participants in the no offloading → no offloading group. These comparisons aimed to explore the impact of prior (*vs*. no

[26] The reported t-tests refers to Tuckey's post-hoc comparisons from the mixed ANOVAs presented in section 3.3.3.2

prior) experience with external aids on perceived task difficulty and effort in the test block where no external aids for offloading strategies were available. The contrast was specified in Linear Mixed Models (LMMs), with group and block as fixed effects and participants' slope included as a random effect. The Satterthwaite method was used to estimate the degrees of freedom.

Perceived Difficulty of Route Planning task

There was a significant main effect of group on perceived task difficulty, $F(4,161) = 8.35$, $p < .001$, $\eta^2_p = 0.17$, but no significant main effect of block, $F(1,161) = 0.21$, $p = .464$, $\eta^2_p = 0.00$. However, the group × block interaction was significant, $F(4,161) = 44.02$, $p < .001$, $\eta^2_p = 0.52$. Post-hoc comparisons revealed that participants in the full offloading → no offloading group perceived the task as more difficult in the test block ($M = 5.41$, $SD = 1.26$) than in the training block ($M = 4.00$, $SD = 0.78$), $t(161) = 7.25$, $p < .001$. Similarly, participants in the partial offloading → no offloading group perceived the task as more difficult in the test block ($M = 5.12$, $SD = 1.36$) compared to the training block ($M = 4.00$, $SD = 1.06$), $t(161) = 5.67$, $p < .001$. Conversely, participants in the no offloading → full offloading group rated the task as less difficult in the test block ($M = 2.82$, $SD = 1.06$) compared to the training block ($M = 4.29$, $SD = 0.94$), $t(161) = -7.55$, $p < .001$. Likewise, participants in the no offloading → partial offloading group perceived the task as easier in the test block ($M = 3.65$, $SD = 1.20$) than in the training block ($M = 4.76$, $SD = 1.10$), $t(161) = -5.74$, $p < .001$. Not surprisingly, participants in the no offloading → no offloading group did not rate the difficulty of the training block ($M = 4.38$, $SD = 1.01$) as significantly different from that of the test block ($M = 4.71$, $SD = 1.13$), $t(161) = 1.27$, $p = .208$.

Tukey’s post-hoc comparisons showed no significant difference in perceived difficulty between the full offloading → no offloading group and the partial offloading → no offloading group, either in the training block, $t(161) = 0.00$, $p = 1.00$, or the test block, $t(161) = 0.98$, $p = .993$. Similarly, no significant difference was found between the no offloading → full offloading and no offloading → partial offloading groups in either the training block, $t(161) = 2.00$, $p = .597$, or the test block, $t(161) = 2.81$, $p = .141$.

Planned comparisons revealed that participants in the full offloading → no offloading and partial offloading → no offloading groups did not rate the task as less difficult than participants in the no offloading → no offloading group in the training block, $t(254.69) = 1.58$, $p = .115$. However, they rated the task as significantly more difficult in the test block, $t(267.90) = 2.48$, $p = .014$.

Since both the groups trained with external aids for offloading strategies (i.e., full offloading → no offloading and partial offloading → no offloading groups) showed a significant increase in perceived difficulty after their removal, and both the groups trained without external aids for offloading strategies (i.e., no offloading → full offloading and no offloading → partial offloading groups) exhibited a significant decrease after their introduction, we computed Δ perceived difficulty scores between the training block and the test block. Similar to the other Δ scores used in this study, a higher Δ perceived difficulty score indicates a greater increase in perceived difficulty after the removal of external aids for offloading strategies for the full offloading → no offloading and partial offloading → no offloading groups, and a greater decrease after their introduction for the no offloading → full offloading and no offloading → partial offloading groups. The one-way ANOVA conducted on Δ perceived difficulty scores did not indicate significant between-group differences, $F(3,131) = 0.84$, $p = .472$, $\eta^2_p = 0.02$.[27]

Perceived Effort of Route Planning task

The main effects of group, $F(4,161) = 9.31$, $p < .001$, $\eta^2_p = 0.19$, and block, $F(1,161) = 7.90$, $p = .006$, $\eta^2_p = 0.05$, on task perceived effort were qualified by a significant group by block interaction, $F(4,161) = 50.81$, $p < .001$, $\eta^2_p = 0.56$. Post-hoc comparisons indicated that participants in the full offloading → no offloading group perceived more effort to complete the test block ($M = 5.74$, $SD = 1.21$) than the training block ($M = 4.26$, $SD = 1.05$), $t(161) = 7.10$, $p < .001$. The same was found for participants in the partial offloading → no offloading group, who reported the task as more effortful in the test block ($M = 5.24$, $SD = 1.15$) compared to the training block ($M = 4.36$, $SD =$

[27] The no offloading → no offloading group was not included in this analysis, as participants in this group did not exhibit significant differences in task perceived difficulty between the training and test blocks.

1.17), $t(161) = 4.18$, $p < .001$. In contrast, participants in the no offloading → full offloading group reported less effort to complete the test block ($M = 2.65$, $SD = 1.32$) compared to the training block ($M = 4.68$, $SD = 1.20$), $t(161) = -9.80$, $p < .001$. Similarly, participants in the no offloading → partial offloading group perceived the task as less effortful in the test block ($M = 3.65$, $SD = 1.28$) than in the training block ($M = 5.06$, $SD = 1.32$), $t(161) = -6.82$, $p < .001$. Participants in the no offloading → no offloading group did not rate the perceived effort for the training block ($M = 4.85$, $SD = 1.35$) as significantly different from that of the test block ($M = 4.74$, $SD = 1.06$), $t(161) = -1.04$, $p = .229$.

Tukey's post-hoc comparisons showed no significant difference in perceived effort between the full offloading → no offloading and partial offloading → no offloading groups in the training block, $t(161) = 0.34$, $p = 1.00$. However, participants in the no offloading → full offloading group reported less effort to complete the test block compared to participants in the no offloading → partial offloading group, $t(161) = -3.40$, $p = .028$.

Planned comparisons revealed that participants in the full offloading → no offloading and partial offloading → no offloading groups rated the training block as significantly more effortful than participants in the no offloading → no offloading group, $t(268.87) = 3.06$, $p = .002$.

Given that both the groups trained with external aids for offloading strategies (i.e., full offloading → no offloading and partial offloading → no offloading groups) showed a significant increase in perceived effort after their removal, and both the groups trained without external aids for offloading strategies (i.e., no offloading → full offloading and no offloading → partial offloading groups) showed a significant decrease in perceived effort after their introduction, we computed Δ perceived effort scores between the training block and the test block. A higher Δ perceived effort score indicates a greater increase in perceived effort after the removal of external aids for offloading strategies for the full offloading → no offloading and partial offloading → no offloading groups, and a greater decrease after their introduction for the no offloading → full offloading and no offloading → partial offloading groups. The one-way ANOVA conducted on Δ perceived effort scores revealed significant between-group differences, $F(3, 131) = 4.69$, $p = .004$, $\eta^2_p = 0.10$.[28] Tukey's post-hoc comparisons indicated that the extent of reduction in perceived effort by participants in the no offloading → full offloading group ($M = 2.03$, $SD = 1.27$) was greater than the extent of increase shown by participants in the partial offloading → no offloading group ($M = 0.88$, $SD = 1.24$), $t(131) = 3.74$, $p = .002$. All other between-groups comparisons were not significant (ts < 2.03, ps > .183). Therefore, when considering the actual extent of increase/decrease in perceived effort between the training and test blocks, the difference between participants in the no offloading → full offloading and no offloading → partial offloading groups was no longer significant. However, a significant difference emerged between participants in the no offloading → full offloading and partial offloading →no offloading groups, indicating that the former participants perceived a greater relief in terms of effort than the burden perceived by the latter ones.

Therefore, consistent with H4a, the removal of external aids for offloading strategies led to a significant increase in perceived task difficulty and effort for participants trained with these aids (full offloading → no offloading and partial offloading → no offloading groups). Conversely, the introduction of aids for offloading strategies decreased these perceptions for participants trained without them (no offloading → full offloading and no offloading → partial offloading groups). Additionally, participants in the full offloading → no offloading and partial offloading → no offloading groups rated the test block as more difficult and effortful than those in the no offloading → no offloading group, possibly highlighting the impact of prior experience with external aids for offloading on perceived difficulty and effort once these aids were removed.

No significant differences in perceived task difficulty were observed between participants provided with external aids for all four offloading strategies and those provided with only two aids, either in the training block or the test block, suggesting that the number of external aids for offloading strategies did not significantly affect perceived task difficulty. However, participants provided with all

[28] The no offloading → no offloading group was not included in this analysis, as participants in this group did not exhibit significant differences in task perceived effort between the training and test blocks.

four aids for offloading strategies during the test block reported a greater decrease in effort compared to the increase in effort experienced by those deprived of two aids for offloading strategies. This provides partial support for H4b.

Table 6

Hypotheses Supported, Partially Supported and Not Supported in Study 3

STUDY 3 - External Aids for Offloading Strategies		
H1: Asymmetric Transfer Effects on Route Planning performance	Extent of performance decline after the removal of offloading greater than the extent of improvement after its introduction.	proportion of maps completed without errors: ***not supported***
		proportional deviation above the shortest route: **supported**
		map completion time: **supported**
H2: Asymmetric Transfer Effects moderated by WM	Less slowdown in solution-typing time after offloading removal for participants with higher verbal WM	**Supported**
H3: More-than-optimal use of offloading strategies	No differences in route planning performance between the full offloading and partial offloading conditions	training block: **supported**
		test block: **supported**
H4a: Transfer Effects on Task Perceptions	Increase in perceived task difficulty and effort after offloading removal; decrease in these perceptions after offloading introduction.	perceived task difficulty: **supported**
		perceived task effort: **supported**
H4b: Differences in Task Perceptions between providing/removing all vs. part of offloading aids	Removing/providing aids for all (vs. part of) the offloading strategies results in a greater burden/relief in perceived task difficulty and effort.	perceived task difficulty: ***not supported***
		perceived task effort: ***partially* supported**

3.3.4. Discussion

Study 3 extended our investigation of transfer effects in route planning by examining the removal or introduction of external aids for cognitive offloading strategies, building on findings from Study 2. Consistent with our first hypothesis (H1), we observed asymmetric transfer effects. Participants trained with external aids for offloading strategies experienced a greater decline in performance when these aids were removed compared to the improvement observed in participants who were provided with such aids after being trained without them. Specifically, participants in the full offloading → no offloading and partial offloading → no offloading groups planned significantly longer routes and required more time to complete them in the test block, whereas participants in the no offloading → full offloading and no offloading → partial offloading groups planned routes faster in the test block than in the training block although they did not exhibit a significant improvement in planning accuracy. These findings mirror the results of Study 2, where participants trained with a pen also experienced a greater performance decline when the offloading opportunity was removed, compared to the improvement observed when it was introduced.

Planned comparisons contrasting the performance of participants in the full offloading → no offloading and partial offloading → no offloading groups with the control group (no offloading → no offloading) in the test block corroborated these findings. Indeed, these comparisons showed that participants who had external aids for offloading strategies removed deviated more from the optimal route and took longer to complete the task compared to the control group, who performed consistently without external aids across both blocks. This suggest that prior experience with external aids for offloading strategies may negatively impact subsequent performance when these aids are no longer available. An alternative explanation could be that participants in the control group (no offloading → no offloading) outperformed those trained with external aids for offloading strategies (full offloading → no offloading and partial offloading → no offloading groups) in the test block due to a learning effect. Specifically, it could be argued that performing both the training and test blocks without external aids may have allowed these participants to develop strategies better suited to completing the task without external supports to offload cognition. However, this explanation seems unlikely, as participants in the no offloading → no offloading group did not show significant improvements between the training and test blocks on either route planning performance measures or task perception measures. This suggests that the control group's performance did not benefit from a learning effect across blocks, further supporting the idea that prior experience with external aids for offloading strategies can negatively affect subsequent performance when those aids are removed.

The moderation analyses provided further insights into transfer effects. In line with H2, the participants trained with external aids who had better verbal working memory exhibited a less pronounced slowdown in solution-typing times when these aids were removed, mirroring the moderation effect of spatial working memory on map completion time observed in Study 2.

These findings are consistent with the idea that abandoning a primarily visual strategy in favour of a strategy mainly based on working memory is more challenging than the opposite shift. Indeed, participants trained with the external aids for offloading strategies struggled to adjust to the route planning task after the removal of these aids. Conversely, participants initially trained without the external aids better adapted to their introduction. These outcomes provide support for the idea that, in our task, shifting from a predominantly visual strategy to a strategy mainly based on working memory is more challenging than the reverse, highlighting that different cognitive consequences on task execution can arise from opposite switches between distinct strategies. Consistent with H3, participants provided with aids for all four offloading strategies did not outperform those with aids for only two strategies in either the training or test block. This finding aligns with results from Study 2, where participants who were trained using the pen employed more offloading strategies during the test block compared to those who were trained without it. However, this greater reliance on offloading strategies did not result in superior performance. Therefore, the outcomes from Study 3 reinforce the idea that, when given the option to offload cognition, individuals may tend to use offloading strategies more than needed for sustaining performance, possibly for reducing cognitive effort. This interpretation is also consistent with prior research on cognitive offloading in other domains (e.g., Gilbert et al., 2023), as well as findings from Study 1.

Participants in the full offloading → no offloading group were more likely to take annotations during the training block compared to those in the no offloading → full offloading group during the test block. This finding generally aligns with the strategy perseveration effect found in Study 2, as well as with previous research on offloading in different cognitive domains (Scarampi & Gilbert, 2020), in that previous experience without the opportunity to offload cognition may hinder the use of cognitive offloading when it is eventually allowed.

Finally, in terms of task perception, the removal of external aids for offloading strategies led to an increase in perceived task difficulty and effort for participants who trained with these aids, while the introduction of external aids had the opposite effect, consistent with H4a. Planned comparisons also showed that participants in the full offloading → no offloading and partial offloading → no offloading groups perceived the task as more difficult and effortful in the test block compared to the no offloading → no offloading group, suggesting that prior (*vs*. no prior) experience with cognitive offloading hindered not only task performance when offloading became unavailable, but affected also task perception.

No significant differences in task perceived difficulty were observed between participants provided with aids for all four offloading strategies and those provided with aids for two only. However, differences did emerge in task perceived effort. Specifically, participants provided with aids for all four offloading strategies perceived a greater reduction in effort compared to the increase in effort perceived by participants deprived of aids for two strategies. This finding resembles the one obtained for task difficulty in Study 2, in which participants allowed to use the pen during planning after training with their hands still perceived a greater reduction in task difficulty compared to participants who were provided with the pen after having been trained without pen use only, providing partial support for H4b.

3.4 Overall Discussion for Study 2 and Study 3

The primary aim of Studies 2 and 3 was to investigate the bidirectional transfer effects of cognitive offloading in route planning, specifically examining how the option to offload or not offload cognition during a route planning task affected subsequent performance when this option was either removed or introduced. By exploring these shifts, we aimed to extend the current understanding of the consequences of cognitive offloading, providing, to our knowledge, the first insights into how these bidirectional transfer effects take place in a complex cognitive task like route planning.

In Study 2, participants freely devised their own offloading strategies using a pen while completing a route planning task. Study 3 embedded visual aids into the maps to support these offloading strategies. This eliminated the need for participants to develop offloading strategies independently.

In agreement with our hypothesis, both studies revealed asymmetric transfer effects on route planning performance. In Study 2, participants trained with pen use experienced a greater decline in performance when the opportunity to use the pen was removed, compared to the improvement observed when it was given to participants initially trained without it. This was seen in both planning accuracy and completion time. Study 3 confirmed and extended these findings, additionally revealing that participants who had been trained with external aids for offloading strategies were less accurate and slower in planning routes once these aids were removed during the testing phase compared to participants in the control group, who completed both the training and the test blocks without external aids.

The asymmetric transfer effects observed in both Study 2 and Study 3 align with our expectation that shifting from a mainly visual strategy to a strategy mainly based on working memory for route planning when offloading is precluded is more challenging than the reverse when offloading is allowed. This expectation was based on findings from Study 1, which suggested that participants with the opportunity to offload cognition would have approached the task primarily as a visual one, effectively solving it through intuitive perceptual routines, similar to those employed in traveling salesman problems (MacGregor & Chu, 2011; MacGregor & Ormerod, 1996). Consistent with this view, when participants had the opportunity to offload cognition, their route planning performance was correlated with visual search abilities in both Study 2 and Study 3 (see, respectively, Appendix B, section 2, and Appendix C, section 1). Conversely, participants trained without the opportunity to offload cognition were expected to adopt an incremental planning strategy mainly based on working memory to plan and retain their routes. Importantly, we assumed that the cognitive effects of shifting between these strategies would differ based on the direction of the transition. Our results supported this assumption, showing that moving from a primarily visual strategy to a strategy primarily based on working memory is indeed more difficult than shifting in the opposite direction.

Our interpretation of the asymmetric transfer effects was further supported by the role of working memory in moderating the challenges of this transition. In Study 2, participants trained with the pen were better able to cope with its removal when they had better spatial working memory. Similarly, in Study 3, participants trained with external aids who had better verbal working memory exhibited a less pronounced slowdown in solution typing times when these aids were removed. Beyond reinforcing the idea that working memory supports task performance when cognitive offloading is unavailable, these results also suggest that individual differences in cognitive abilities may buffer the effects of strategy shifts by facilitating the transition to strategies requiring internal cognitive resources.

Study 2 and Study 3 build on prior research showing that cognitive offloading can impair subsequent unaided performance, particularly highlighting the negative effects of offloading information on subsequent memory for that information (e.g., Fenech, 2010; Henkel, 2014; Sparrow, 2011). Our findings extend this body of work by further demonstrating the potential detrimental effects of prior cognitive offloading on subsequent unaided performance within the same route planning task, but also showing that participants can quickly learn how to offload cognition when allowed to do so, despite a previous experience with the task when offloading was precluded. Thus, it seems that the specific interaction between the strategies used by participants over time matters, with strategy selection triggered by the availability/unavailability of the opportunity to offload cognition. Our findings add to this literature by providing evidence that the ease or difficulty of adapting from one strategy to another is contingent on the specific cognitive features of both the initial and final strategies.

These results offer a more complex and articulate picture that the ones that can be reached starting from the literature on transfer effects in problem-solving that did not consider the effects of the opportunity to offloading cognition, such as the studies on negative transfer of learned strategies (Bilalić et al., 2008; Bilalić & McLeod, 2014; Luchins & Luchins, 1950; Ohlsson, 2011), and the studies on the positive consequences of learning to perform a task with a planful strategy triggered by high operator cost (O'Hara & Payne, 1998).

Consistent with our hypotheses, Study 2 also provided evidence for a strategy perseveration effect. In the test block allowing pen use, participants trained with the pen employed more offloading strategies than those trained without it. Specifically, participants who were trained with the pen marked ordering rules more frequently and kept an external track of their plan, whereas no significant differences emerged between groups in marking the starting and arrival locations or intermediate locations. This increased reliance on offloading strategies, however, did not translate also into better performance during the test block, suggesting that participants trained with the pen may have engaged in a more than optimal use of offloading strategies. Prior research on cognitive offloading in other domains has similarly shown a bias towards adopting external strategies, even when this does not enhance task performance (e.g., Gilbert et al., 2020; Gilbert et al., 2023; see also Study 1). To investigate this potential over-reliance further, Study 3 directly assessed whether providing external aids for all four offloading strategies (i.e., full offloading condition) resulted in better route planning performance than providing aids for two strategies only (i.e., partial offloading condition, where only the two strategies with no between-group differences were aided into the maps). The results revealed no significant performance differences between these offloading conditions in either the training or test blocks. Therefore, the findings of Study 3, in conjunction with those of Study 2, support the idea that individuals may show a bias toward cognitive offloading while planning routes, resorting to it more than necessary to support performance. These results extend previous findings on cognitive offloading

(Gilbert et al., 2020; Gilbert et al., 2023), specifically by showing that an over-reliance on offloading can also occur in the domain of route planning. However, unlike prior research, which focused on over-reliance in terms of frequency of use a single offloading strategy, our studies highlighted this over-reliance across both the frequency and the variety of offloading strategies employed.

Results from Study 2 also suggest that the perseverative and more than optimal use of offloading strategies among participants trained with the pen may have contributed to the observed asymmetric transfer effects. Specifically, our mediation analysis showed that the greater the reliance on offloading strategies by participants who were trained with the pen, the more they struggled with adapting to a working memory-based strategy when pen use was removed. This may indicate that the offloading strategies, which were consistently employed across the training and the training-consistent test blocks ($r(69) = .84$, $p < .001$; see also Appendix B, section 3), potentially interfered with participants' ability to adapt to the route planning strategy adopted when the option to use the pen was no longer available. Supporting this interpretation, Study 3 revealed that participants trained with external aids for offloading strategies were less accurate and slower in the test block, when these aids were removed, compared to participants in the control group, who completed both the training and the test blocks without any external aids. This may indicate that, also in Study 3, the availability of external aids for offloading may have interfered with the adoption of the more cognitively demanding and primarily working memory-based strategy. It is worth noting that, in Study 3, we also observed no significant differences in the performance decline between participants deprived of all versus only some aids for offloading strategies. This finding diverges somewhat from the results of the mediation analysis in Study 2, suggesting that, while prior experience with offloading aids hindered the transition to a working memory-based strategy, the extent of this interference was not influenced by the number of offloading strategies supported. However, the mediation analysis from Study 2 also demonstrated a direct effect of the initial training condition on the extent of changes in performance during the test phase. This implies that the initial training condition influenced subsequent performance not only through the number of offloading strategies used, but also directly, likely establishing a distinct mental framework akin to a “special kind of mental set” for the use of at least some of the offloading strategies (Luchins & Luchins, 1950).

That is, participants who were trained with offloading option struggled to abandon their previous strategy, which was particularly challenging to replace due to its reliance on intuitive visual routines (see, e.g., MacGregor & Chu, 2011). This difficulty in shifting from a predominantly visual strategy would appear even when not all offloading strategies were devised (Study 2) or externally supported (Study 3). In contrast, participants trained without the option to offload cognition might have adapted more readily when offloading was introduced, as this allowed them to switch to a visual strategy that was more intuitive, less cognitively demanding, and thus easier to be adopted.

The ease with which participants trained without the option to offloading cognition adapted to a more intuitive, predominantly visual strategy might also be generally related to the study by O'Hara and Payne (1998). They found that participants initially trained under a high-cost operator condition in a problem-solving task, compared to participants who initially trained under a low-cost operator condition, adapted more readily and performed better in a subsequent common low-cost condition. The authors attributed this to the high-cost training fostering a more structured and planful approach to the task. However, in the test block allowing pen use in Study 2, we did not find that participants trained without the option to offload cognition outperformed those trained with it, a result that is inconsistent with the explanation provided by O’Hara and Payne.

Study 2 and Study 3 also investigated the effects of removing or providing the possibility to offload cognition on task perception in terms of perceived difficulty and effort. They consistently revealed that removing the possibility to offload cognition for participants who had initially trained with this option led to increased perceptions of task difficulty and effort. Conversely, introducing the possibility to offload cognition for participants trained without this option decreased these perceptions. However, unlike performance measures, these effects were not asymmetric, as the extent of perceived burden following the removal of cognitive offloading was not significantly different to the extent of relief that followed its introduction. Interestingly, findings of Study 2 revealed that participants trained without both pen and hand use experienced a greater reduction in task perceived difficulty when later provided with the pen, compared to those trained only without the pen. Notably, this occurred even though no differences in route planning performance were observed between these two groups during either the training or test blocks. A possible explanation for this might be that the restriction on hand use, while not significantly affecting performance, heightened participants' subjective experience of task difficulty. When this restriction was removed, participants may have experienced a greater sense of relief, leading to a more pronounced decrease in perceived difficulty, compared to those restricted only from pen use. Building on these findings, in Study 3, we hypothesized that removing or providing all four offloading strategies, as opposed to only two, would similarly lead to a greater sense of relief or burden, resulting in more pronounced changes in perceived task difficulty and/or effort. We found that participants provided with aids for all four offloading strategies reported a greater reduction in perceived effort, compared to the increase in effort reported by participants deprived of aids for two strategies, which partially supported our hypothesis.

Overall, findings on the measures of task perception in Studies 2 and 3 appear to align with previous accounts in the literature on cognitive offloading, which suggest that the avoidance of cognitive effort may be a key factor driving the tendency to offload cognition (Sachdeva & Gilbert, 2020; Gilbert et al., 2023; see also Study 1). More broadly, findings from Study 2 and Study 3 indicate that a greater reliance on cognitive offloading, even when it does not improve task performance, could still offer benefits in terms of perceived task difficulty or effort. Thus, it is conceivable that the more-than-optimal use of offloading might be explained by participants' need to reduce perceived effort or difficulty in a complex task, rather than by their need to enhance performance.

In conclusion, Studies 2 and 3 provide evidence for asymmetric transfer effects when shifting between offloading and non-offloading conditions in a route planning task, with performance declining more when offloading was removed than improving when it was introduced. These effects highlight the challenge of shifting from an external, primarily visual strategy, to an internal strategy, primarily based on working memory. Study 2 also revealed that participants trained with the possibility to offload cognition tended to over-rely on offloading strategies, even when such reliance did not improve performance. Study 3 reinforced this finding, as giving full versus partial aids for offloading strategies did not enhance performance, providing further evidence for a more-than-optimal use of cognitive offloading also in the domain of route planning. Additionally, both studies offered partial support for the idea that individuals may overuse cognitive offloading to alleviate perceived task difficulty and effort.

A main limitation of the studies presented so far is that participants were presented with artificially created maps, which may limit the generalizability of our findings. Study 4 represents an initial attempt to address this limitation by investigating the effects of cognitive offloading on route planning in a real-world map of a suburb of Berlin, potentially broadening the applied implications of our findings.

STUDY 4.

COGNITIVE OFFLOADING AND ROUTE PLANNING IN A REAL-WORLD MAP

4.1 Introduction

Studies 1, 2, and 3 investigated cognitive offloading in a complex task requiring participants to plan routes on maps under ordering constraints. Study 1 identified the offloading strategies that individuals spontaneously devised to offload cognition during route planning and demonstrated their effectiveness in supporting performance, particularly when the task was more challenging. Study 2 observed the use of the same spontaneous offloading strategies and revealed that removing or introducing the opportunity to use them produced asymmetric transfer effects on route planning performance. Study 3 extended these findings by showing that, when these offloading strategies were supported by external aids embedded into the maps, they continued to produce asymmetric transfer effects when the aids were introduced or removed. However, all three studies used artificially created maps, which may limit the generalizability of the findings. Study 4 aimed to address this limitation by using more ecologically valid materials, specifically presenting participants with a map of a real-world location.

4.1.1 Main Aims and Hypothesis

The main aim of Study 4 was to assess whether the external aids for offloading strategies tested in previous studies would remain effective when applied to a real-world map, thus allowing us to evaluate the generalizability of part of our earlier findings and explore their potential real-world applications. To this end, participants were presented with a map of part of a real city and asked to plan the shortest route according to specified constraints under one of two between-participants conditions. In the no offloading group, participants were provided only with the map and task instructions, like in the no offloading conditions of Study 3. In the offloading group, the map also embedded external aids for offloading strategies, again similarly to Study 3. In line with our previous findings, we hypothesized that participants in the offloading group would demonstrate better route planning performance than those in the no offloading group (H1).

Additionally, we anticipated replicating our previous findings on task perceived difficulty and effort, predicting that participants in the offloading condition would report lower levels of perceived difficulty and effort compared to participants in the no offloading condition (H2).

Finally, we aimed to assess whether the availability of external visual aids for offloading cognition would trigger the cognitive strategies underlying route planning in the manner suggested by our earlier studies. Specifically, participants in the no offloading condition were expected to rely on an incremental strategy, primarily based on working memory to plan the route. Without external aids, they would need to progressively add the to-be-visited locations to the planned route while mentally rehearsing the location names in sequence. This working memory-based strategy was hypothesized to lead to a deeper encoding of location names, resulting in better *verbal* recall of the planned route for participants in the no-offloading condition, as compared to those in the offloading condition (H3a). In contrast, participants in the offloading condition were expected to rely on a predominantly visual strategy to support their route planning. By offloading the need to remember the locations and planning constraints onto visual aids, these participants could focus more on the visual-spatial aspects of the map, such as organizing locations based on their positions. As a result, we hypothesized that participants in the offloading condition, compared to those in the no-offloading condition, would exhibit better recall of the spatial positions of the planned route (H3b). If supported, these hypotheses would provide further evidence for the cognitive mechanisms involved in the strategy shifts investigated in Studies 2 and 3.

4.2. Method

4.2.1 Participants

An a priori power analysis determined that at least 90 participants were needed to detect medium effect sizes for this experimental design ($f = 0.25$, $\alpha = 0.05$, $1 - \beta = 0.80$). A total of 101 undergraduate students took part in the study. Eight participants, four from the offloading condition, and four from the no offloading condition, exited the experiment after completing the practice trials but before completing the test trial and the subsequent part of the experiment. The final sample consisted of 93 participants (82.20% female, age: $M = 22.88$, $SD = 5.46$).[29] The research protocol was approved by the Ethical Committee of the University of Trieste, and informed consent was obtained at the beginning of the online experimental session.

4.2.2 Experimental Design

The study employed a 2 (group: offloading *vs*. no offloading) × 2 (recall task: verbal *vs*. visual) mixed design. The availability of external aids for offloading strategies was manipulated between participants: those in the offloading condition were presented with a map that incorporated external aids for offloading strategies, while those in the no offloading condition were provided with a map without these aids.

4.2.3 Procedure

The entire experimental procedure was designed and administered using the Qualtrics XM software. At the start of the experimental session, participants read the study information and provided informed consent. They were then randomly assigned to one of two groups:

1) Offloading, in which external aids for three offloading strategies were provided in both the practice and the test maps (n = 46);

2) No offloading, in which no aids for offloading strategies were embedded into the practice and the test maps (n = 47).

Participants first read the instructions for the route planning task and completed two practice maps. Afterwards, they moved to the test map. Upon completing the test map, they rated the difficulty and effort required for planning the route.

[29] A total of 3 participants, one from the offloading condition, and two from the no offloading condition, preferred not to indicate their age.

Next, participants were asked to recall the names and the spatial positions of the places in the planned route, with the order of these tasks counterbalanced across participants. After completing both recall tasks, participants undertook an additional recall task, in which they recalled the route again, this time specifying both the names of the locations and their positions on the map in the planned sequence.

Following the recall tasks, participants were asked whether they had previously participated in a study involving route planning on maps, and whether they used any further external support during the previous tasks (e.g., paper and pen, smartphone, notes on the computer). Finally, the participants answered demographic questions.

The entire procedure took approximately 20-30 minutes. After the data collection, a debriefing session was conducted, during which participants were fully informed about the aims of the study and provided with additional details.

4.2.4 Materials

4.2.4.1 Route Planning Task

A computerized version of the route planning task was created using the Qualtrics XM software. The maps used in this study were taken from an online repository (https://stock.adobe.com/) and specifically selected for their vector-based design without verbal labels. Each participant was presented with three maps: two for practice trials, and one for the test trial.

Both the practice maps and the test map included a starting point, an endpoint, eight intermediate locations to be visited, two ordering rules (which constrained the order of visit for two pairs of intermediate locations), and ten distractor locations. The method for creating the map directions followed a procedure similar to the one used in our previous studies. Specifically, an 8×5 rectangular matrix was superimposed over each map, from which 20 out of 40 points were randomly selected. Once the matrix was removed, the plausibility of the positioning of each selected point was verified to ensure it fell on a street within the map. If a point did not correspond to a street, it was moved to the nearest street. Subsequently, a starting point, an endpoint, and eight intermediate points to be visited were randomly selected. Additionally, two ordering constraints were introduced by randomly selecting two pairs of intermediate locations (excluding the starting and endpoints) and determining their order of visit.

Unlike previous studies, where locations were labelled using letters from the Latin alphabet, this study provided more realistic labels for both target and distractor locations. One practice map was named “City of Trees” (*Città degli Alberi*), with all locations labelled using the names of 20 among the most known tree species in Italy. The other practice map was called “City of Friends” (*Città degli Amici*), with locations labelled using 20 among the most common Italian first names (10 female and 10 male). These specific labels (tree species and first names) were chosen to prevent interference from the location names of the practice maps to the actual test map, which used common names of urban locations. Specifically, by using distinct semantic categories for labelling locations of each map, we aimed to minimize the possibility that the names encoded during the practice session would interfere with the recall of the names of the test map (Del Missier et al., 2018; Loess, 1968). The presentation order of the two practice maps was counterbalanced across participants to control for potential order effects.

For the test trial, the map represented an area of Berlin. To minimize influences from potential participants' familiarity with the city, a peripheral northwest area was chosen, and the map was referred to as “Vacation City” (*Città delle Vacanze*). Furthermore, we named the locations using realistic but fictional labels representing common urban places, such as Station, Restaurant, Museum, Post Office, University, and so on (see Figure 20). The 20 labels used in each map were randomly assigned to the 20 points for that map.

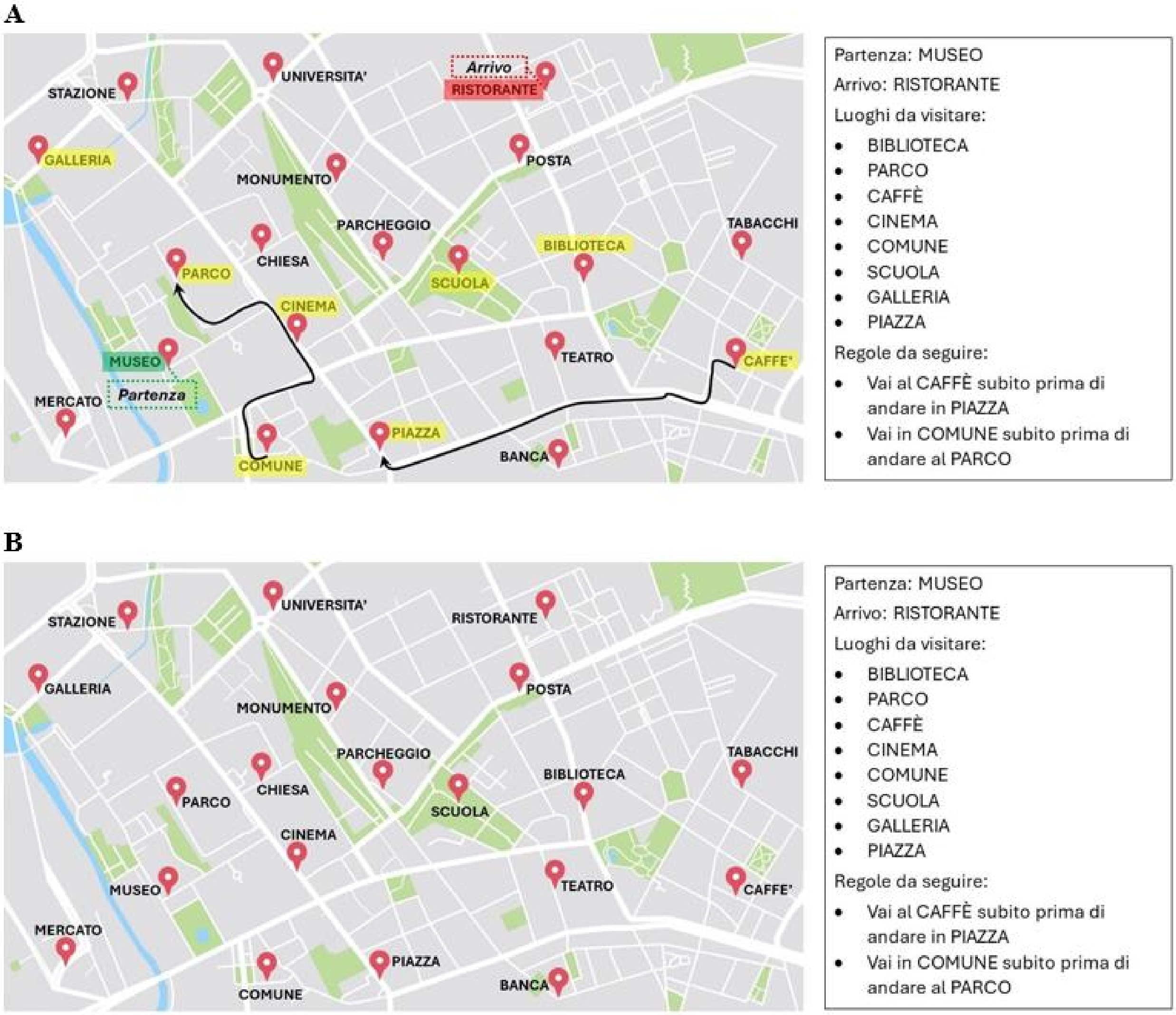
Participants first read the instructions for the route planning task and completed the practice maps under their assigned condition. They were instructed to plan the shortest route as quickly and accurately as possible. After determining the full route, participants clicked a "Next" (*Avanti*) button to access the response field below the map, where they typed the names of the locations in the exact order they planned to visit them. The instructions were consistent across the two conditions.

Similar to the full offloading condition in Study 3, participants in the offloading group of Study 4 were shown an example map with specific directions (e.g., starting and ending locations, intermediate locations to be visited, and ordering rules) alongside visually marked starting/arrival points and intermediate locations. Additionally, locations constrained by ordering rules were connected by arrows. However, unlike Study 3, no textbox for tracking the plan was provided. This decision was based on findings from Study 3, in which not all participants used the textbox to take notes during route planning in the full offloading condition. In the no offloading group, participants were shown an example map with the directions on the right, but without any external visual aids, mirroring the no offloading condition of Study 3.

After reading the instructions, participants completed two practice trials with two different maps. If they made errors (e.g., incorrect starting/arrival points or violation of ordering rules) on the first practice trial, their typed route was displayed along with feedback highlighting the error(s), and they were asked to repeat the trial. If errors persisted after a second attempt, participants were shown an example of a correct route that met all task criteria. This same procedure was followed for the second practice trial as well. The decision to include two practice maps, rather than just one as in previous studies, was based on two considerations. First, the maps in this study were visually more complex than those previously used, potentially requiring additional practice to ensure that participants fully understood the task requirements. Second, given that only one map was used in the test trial, it was essential to ensure that participants had fully understood the task before completing it.

Figure 20

Map Used in the Route Planning Task for the Offloading (A) and No Offloading (B) Groups



Note. In Panel A, the starting place (*Partenza*) and arrival place (*Arrivo*) are verbally labeled and visually highlighted in green and red, respectively, within the map. Locations to be visited (*Luoghi da visitare*) are highlighted in yellow, and ordering rules (*Regole da seguire*) are represented as directed black arrows. In Panel B, no visual aids for offloading strategies are provided.

Upon completing the practice maps, participants were informed they would be presented with a new map for the actual route planning task. Before proceeding to the test map, they were asked to forget the previous maps and routes to better focus on the upcoming map. This "directed forgetting" instruction (e.g., Geiselman et al., 1983) was introduced as an additional precaution to reduce interference from practice maps.

The scoring procedure mirrored that used in the previous studies. First, we measured the proportion of maps completed without errors (i.e., correct starting/arrival locations, inclusion of all intermediate locations, and adherence to ordering rules) in each group. For error-free maps, we then assessed planning accuracy by identifying the shortest route identified by participants and calculating the proportional deviation above this route using the formula: *proportional deviation* = (actual route length - shortest route length) / shortest route length. This metric reflects the quality of the planned route, with higher values indicating greater deviation from the shortest route.

For planning speed, we recorded the time taken to complete each error-free map (the interval between the map presentation and the participant confirming the typed route). Additionally, we measured the time spent planning the route (the interval between the map presentation and the participant accessing the response frame) and the time spent typing the route (the interval between accessing the response frame and confirming the typed route)

4.2.4.2 Perceived Difficulty and Effort of the Route Planning task

Perceived Difficulty and Effort. After completing the test map, participants rated how difficult they found the task and how much effort they made in planning the route. These ratings were given on a 7-point scale, ranging from "very little" to "very much."

4.2.4.3 Recall of the Planned Route

Verbal Recall of Location Names. Participants were presented with a response frame similar to the one below the task map, but without the map itself. This frame displayed the numbers 1 to 10 horizontally, with a space beneath each number. By clicking on each space, a drop-down menu appeared, listing all the names of map locations, including distractors, in alphabetical order. Participants were asked to indicate the route they had identified earlier by selecting, for each number, the name of the location corresponding to that position in their planned route. The software did not force participants to respond, but the instructions encouraged them to provide an answer even if they were uncertain. We calculated the number of locations correctly recognized, regardless of their order, as well as the names recalled in their correct position in the planned route. Additionally, the time taken to complete the verbal recall task was recorded.

Visual Recall of Location Positions. Participants were shown the actual task map again. However, the map now displayed Latin alphabet letters in place of the original location labels. As in the task map, a response frame appeared below, with numbers 1 to 10 horizontally, each accompanied by a clickable space. Upon clicking each space, a drop-down menu listing all the letters used to denote map spatial locations, including distractors, appeared in alphabetical order. Participants were asked to reconstruct their previously identified route by selecting, for each number, the letter corresponding to the spatial location in that ordinal position within their planned route. While responses were not mandatory, participants were encouraged to complete the task even if uncertain. Similar to the verbal recall task, we calculated the number of spatial locations correctly recalled, regardless of order, as well as locations recalled in their correct position in the planned route. The time taken to complete the visual recall task was also recorded.

Paired Recall. For the paired recall task, participants were again presented with the actual task map, where locations were labelled with Latin alphabet letters in place of their original names. Below the map, two response frames labelled "Letters" and "Locations" were displayed, each featuring numbers from 1 to 10 arranged horizontally with a space beneath each number. The "Letters" frame was similar to the one used in the Visual Recall task. Participants clicked on the space under each number and selected from a drop-down menu the letter corresponding to the spatial location in that position in their planned route. The "Locations" frame mirrored the one used in the Verbal Recall task. For each number, participants selected from a drop-down menu the name of the location that occupied that position in their planned sequence.

Unlike the previous recall tasks, however, the drop-down menus in both the “Letters” and “Locations” frames displayed only the locations that were required to be included in the route, but not the distractor locations. Similarly, the map displayed only the letters corresponding to the locations that participants had to visit, with no distractors. Participants were encouraged by the instructions to complete both parts of the task even if they were uncertain, but responses were not mandatory.

We assessed the accuracy of the sequential recall in relation to the initially typed route for both the spatial positioning of the locations on the map (letters) and their names (original labels). Additionally, we evaluated the accuracy of associations between locations on the map and their names, as well as the correct sequential order of the correctly associated pairs. As with previous recall tasks, we also recorded the time taken to complete the paired recall task.

4.2.4.4 Post-Task Checks for Potential Confounds

After completing the recall tasks, participants answered two additional questions to control for potential confounds. The first question asked if they had previously participated in a study involving route planning on maps, allowing us to control for the potential influence of prior experience.

The second question probed whether participants used any additional supports, such as paper and pencil, smartphone, or computer annotations, during the tasks completed during the experiment. Responses for both questions were collected on a dichotomous yes/no scale.

4.3 Results

4.3.1 Preliminary Analyses

For each performance variable of the test trial, as well as for measures from the recall tasks, data were trimmed within groups. Values that were more than 3 *SD*s above or below the mean were replaced with the mean ± 3 *SD*s (for a similar procedure, see Study 1, Study 2, Study 3; Del Missier et al., 2021; Miyake et al., 2000). Descriptive statistics for all measures of the route planning task by group are presented in Table 7, while Table 8 provides descriptive statistics for the main recall task measures by group.[30]

[30] Data trimming resulted in the following percentages of value substitution. For proportional deviation above the shortest route: 2.63% in the offloading group and 0.00% in the no offloading group. For map completion time: 2.63% in the offloading group and 0.00% in the no offloading group. For route planning time: 2.63% in the offloading group and 0.00% in the no offloading group. For solution-typing time: 5.26% in the offloading group and 2.63% in the no offloading group. For time spent on verbal recall of the route: 2.17% in the offloading group and 0.00% in the no offloading group. For

Table 7

Descriptive Statistics of Route Planning Performance Measures for the Test Map

	Map Completed Without Errors % (N)	Solution-Typing Time M (SD)	Map Completion Time M (SD)	Proportional Deviation Above the Shortest Route M (SD)
Offloading	82.61% (38)	62.21 (45.72)	125.35 (75.66)	0.06 (0.04)
No Offloading	80.85% (38)	82.97 (69.31)	253.83 (101.41)	0.05 (0.04)

Note. Time measures are indicated in seconds.

Table 8

Descriptive Statistics of Recall Tasks

	Verbal recall			Visual recall			Paired recall		
	Time M (SD)	Correct names M (SD)	Correct name order M (SD)	Time M (SD)	Correct locations M (SD)	Correct location order M (SD)	Time M (SD)	Correct pairs M (SD)	Correct pair order M (SD)
Offloading	108.99 (62.45)	8.12 (1.88)	5.28 (3.22)	123.30 (50.76)	6.30 (1.35)	5.26 (2.40)	179.76 (74.38)	4.85 (3.20)	4.41 (3.14)
No Offloading	96.59 (45.31)	8.96 (1.65)	6.89 (3.07)	117.61 (55.78)	7.00 (1.78)	5.91 (3.25)	167.13 (62.76)	6.45 (3.26)	5.85 (3.54)

Note. Time measures are indicated in seconds.

In total, 29.03% of participants reported using supports during the experiment, and 58.06% indicated prior participation in a study requiring route planning on maps. A chi-square test of independence was conducted to examine whether the offloading and no offloading groups differed in their use of additional supports. The test revealed a significant association between group and support use, $\chi^2(1, N = 93) = 5.99$, $p = .014$, with a higher percentage of participants in the no offloading group (40.43%) reporting the use of additional supports compared to the offloading group (17.39%). Given that participants in the offloading and no offloading groups differed in their use of additional supports during the experiment, all subsequent analyses were conducted controlling for external support use.

accuracy in verbal recall of location names: 2.17% in the offloading group and 0.00% in the no offloading group. For time spent on visual recall of the route: 2.17% in the offloading group and 2.13% in the no offloading group. For accuracy in visual recall of location positions: 4.35% in the offloading group and 2.13% in the no offloading group. For time spent recalling the route in the paired task (both verbal and visual): 2.17% in the offloading group and 2.13% in the no offloading group. No substitutions were made for other measures.

A second chi-square test was conducted to assess whether the groups differed in their prior participation in route planning studies. No significant association was found, $\chi^2(1, N = 93) = 3.25$, $p = .071$. Additionally, the two experimental groups did not differ significantly in gender distribution, $\chi^2(1, N = 93) = 0.13$, $p = .720$.

4.3.2 Route Planning Performance on the Test Map

To test the hypothesis that participants in the offloading group would exhibit better route planning performance compared to participants in the no offloading group (H1), a series of regression analyses were performed on route planning task measures.

To investigate whether groups differed in the proportion of participants completing the test map without errors, we conducted a hierarchical binomial logistic regression. In the first block, group was entered as the predictor. In the second block, use of additional supports was added, and in the third block, the interaction between group and external supports was included.

For continuous measures of route planning performance (i.e., proportional deviation above the shortest route, map completion time, route planning time, and solution-typing time), linear multiple regressions were performed. Given that group and additional support use were correlated, and unbalanced numbers of participants were present across different combinations of the two factors, regression models were chosen over ANOVA due to their greater flexibility in estimating effects when factors are correlated, as ANOVA requires orthogonality between factors (Cohen et al., 2003). In multiple linear regression analyses, group was entered in the first block as the predictor, additional support use was added in the second block, and their interaction was included in the third block[31].

Participants completing the test map without errors

None of the factors entered in the binomial logistic regression were significant predictors of the proportion of participants completing the test map without errors. Specifically, the model with only

[31] When previous participation in a study involving route planning on maps was entered as a predictor in the second block of the regression analyses, results for route planning performance measures did not change, except for some results regarding map completion time (see note 32) and route planning time (see note 33), which did not affect the main conclusions on the differences between the offloading and no offloading groups.

group as a predictor was not significant, $\chi^2(1, N = 93) = 0.05$, $p = .826$, $R^2_{McF} = .001$. Similarly, the model including additional support use, $\chi^2(2, N = 93) = 1.43$, $p = .490$, $R^2_{McF} = .016$, and the model adding also the group by external support use interaction, $\chi^2(3, N = 93) = 2.45$, $p = .484$, $R^2_{McF} = .028$, were also not significant.

Therefore, regardless of the use of additional supports, the proportion of participants who completed the map without errors in the offloading group (82.61%) was similar to the proportion in the no offloading group (80.43%).

Proportional Deviation above the Shortest Route

Regarding the proportional deviation above the shortest route, none of the regression models reached statistical significance, Fs < 0.64, R^2s < .02, ps > .430. Regardless of additional support use, participants in the offloading group ($M = 0.06$, $SD = 0.04$) deviated from the shortest route to a similar extent as those in the no offloading group ($M = 0.05$, $SD = 0.04$). Similarly, regardless of group, participants who used additional supports ($M = 0.05$, $SD = 0.04$) deviated from the shortest route comparably to those who did not use such supports ($M = 0.06$, $SD = 0.04$). Descriptive statistics for the proportional deviation above the shortest route for each group by additional support use are provided in Table 9.

These findings indicate that participants provided with visual aids into the map (i.e., offloading group) did not plan shorter routes compared to those without visual aids (i.e., no offloading group), not supporting H1. Similarly, the use of additional forms of support during the experiment did not enhance route planning accuracy.

Table 9

Proportional Deviation Above the Shortest Route for Group by External Support

Offloading		**No Offloading**	
Supports Not Used	**Supports Used**	**Supports Not Used**	**Supports Used**
M (SD)	M (SD)	M (SD)	M (SD)
0.06 (0.04)	0.06 (0.02)	0.05 (0.03)	0.05 (0.04)

Map Completion Time

Regarding map completion time, the starting model including only group explained a significant amount of variance, $R^2 = .35$, $F(1,74) = 39.18$, $p < .001$. Including additional support use to the second model significantly increased the explained variance, $\Delta R^2 = .05$, $F(1,73) = 6.28$, $p = .014$. The final model, which also included the interaction between group and additional support use, further increased the explained variance on map completion time, $\Delta R^2 = .03$, $F(1,72) = 3.96$, $p = .050$ ($R^2 = .43$, $F(3,72) = 18.06$, $p < .001$). Although this increase was only marginally significant, given its proximity to statistical significance and the possibility of limited statistical power, the model was retained for further exploration.

The interaction between group (0 = no offloading, 1 = offloading) and support use (0 = supports not used, 1 = supports used) was marginally significant, $\beta = -.90$, $t(72) = -1.99$, $p = .050$, VIF = 4.10, tolerance = 0.24, suggesting that the effect of group on map completion time depended on whether participants used additional support. Systematic changes in the reference levels of group and external support use indicated that both when participants used and did not use additional supports, those in the offloading group completed the map significantly faster than those in the no offloading group (supports not used: offloading group, $M = 126.33$ s, $SD = 80.88$ s, no offloading group, $M = 217.92$ s, $SD = 92.34$ s, $\beta = -.83$, $t(72) = -3.98$, $p < .001$, VIF = 1.40, tolerance = 0.71; supports used: offloading group, $M = 118.87$ s, $SD = 23.70$ s, no offloading group, $M = 308.90$ s, $SD = 91.72$ s, $\beta = -1.73$, $t(72) = -4.34$, $p < .001$). Among participants in the offloading group, those who used additional supports were not significantly faster than those who did not, $\beta = -.07$, $t(72) = -0,18$, $p = .855$, VIF = 3.39, tolerance = 0.29. However, among participants in the no offloading group, those who used additional supports were significantly slower compared to those who did not, $\beta = .83$, $t(72) = -3.24$, $p = .002$. Figure 21 provides a graphical representation of these findings.

Therefore, regardless of the use of additional supports, participants who were provided with visual aids for offloading strategies (i.e., offloading group) completed the map faster than those without visual aids (i.e., no offloading group), supporting H1. Furthermore, participants in the no offloading group took even longer to complete the map when using (*vs*. not using) external supports[32].

Route Planning Time

The regression model including only group explained a significant amount of variance in route planning time, R^2 = .29, $F(1,74) = 30.52$, p <.001, and the inclusion of additional support use significantly increased the explained variance, $\Delta R^2 = .19$, $F(1,73) = 27.40$, $p < .001$. The final inclusion of the group by support use interaction further increased the explained variance in route planning time, $\Delta R^2 = .06$, $F(1,72) = 8.87$, $p = .004$ (R^2=.54, $F(3,72) = 28.36$, $p < .00$).

In the final regression model, the group (0 = no offloading, 1 = offloading) by support use (0 = supports not used, 1 = supports used) interaction was significant, β = -1.20, $t(72) = -2.98$, p = .004, VIF = 6.36, tolerance = 0.16. Further probing of the interaction effect by systematically changing reference levels revealed that, among participants who reported not using additional supports during the experiment, those in the offloading group (M = 55.60 s, SD = 56.47 s) spent significantly less time in planning the route than those in the no offloading group (M = 109.89 s, SD = 72.74 s), β = -.54, $t(72) = -2.88$, $p = .005$, VIF = 5.07, tolerance = 0.20. Among participants who reported using additional supports, those in the offloading group (M = 78.43 s, SD = 19.74 s) again planned the route significantly faster than those in the no offloading group (M = 253.43, SD = 95.44), β = -1.74, $t(72) = -4.88$, p < .001.

In the offloading group, using additional supports did not significantly change route planning time compared to not using them, β = 0.23, $t(72) = 0.68$, $p = .496$, VIF = 1.62, tolerance = 0.62. In the

[32] Regarding map completion time, adding previous participation in a study involving route planning as a predictor in the second block partially changed the regression results. Specifically, previous participation in a route planning study (0 = no previous participation, 1 = previous participation) significantly predicted map completion time, β = -.49, $t(71) = -2.75$, p = .007, VIF = 1.09, tolerance = 0.92. Participants with prior experience in a route planning study completed the map faster (M = 151.64 s, SD = 76.04 s) than those without prior experience (M = 236.47, SD = 127.17). Additionally, the interaction between group (0 = no offloading, 1 = offloading) and additional support use (0 = supports not used, 1 = supports used) only approached significance, β = -.75, $t(71) = -1.72$, $p = .090$, VIF = 4.16, tolerance = 0.24. Consistently with H1, Group remained a significant predictor of map completion time, β = -.82, $t(71) = -4.09$, $p < .001$, VIF = 1.40, tolerance = 0.71. Additional support use also remained a significant predictor, β = .66, $t(71) = 2.62$, $p = .011$, VIF = 1.71, tolerance = 0.58.

no offloading group, those who used additional supports took significantly longer to complete the task compared to those who did not, β = 1.43, $t(72) = 6.22$, $p < .001$.

Overall, participants in the offloading group consistently completed the route planning task faster than those in the no offloading group, regardless of additional support use, consistent with H1. Specifically, in the absence of additional support use, participants in the offloading condition took nearly half the time to plan the route compared to those in the no offloading condition. When additional supports were used, participants in the offloading group took less than one-third of the time compared to those in the no offloading group. Additionally, using supports significantly slowed down participants in the no offloading group, while it did not significantly affect participants in the offloading group[33].

Solution-typing Time

The model including only group did not explain a significant amount of variance in solution-typing time, $R^2 = .03$, $F(1,74) = 2.37$, $p = .128$ ($R^2 = .11$, $F(2,73) = 4.42$, $p = .015$). However, the inclusion of additional support use significantly increased the explained variance, $\Delta R^2 = .08$, $F(1,73) = 6.30$, $p = .014$. The final addition of the group by additional support use interaction did not further increase the explained variance, $\Delta R^2 = .00$, $F(1,72) = 0.38$, $p = .542$.

In the retained model, group (0 = no offloading, 1 = offloading) was a significant predictor of solution-typing time, β = -.52, $t(72) = -2.27$, $p = .026$, VIF = 1.10, tolerance = 0.91, indicating that participants in the no offloading group, $M = 82.97$ s, $SD = 69.32$ s, took significantly longer to indicate their solution compared to those in the offloading group, $M = 62.21$ s, $SD = 45.72$. Also support use was a significant predictor (0 = supports not used, 1 = supports used), β= -0.66, $t(72) = -2.51$, $p = .014$, VIF = 1.10, tolerance = 0.91, with participants who did not use additional supports, $M = 80.05$ s, SD =

[33] Regarding route planning time, adding previous participation in a study involving route planning as a predictor in the second block did not significantly change the regression analysis results. However, prior participation (0 = no previous participation, 1 = previous participation) significantly predicted route planning time, β = -.33, $t(71) = -2.03$, $p = .046$, VIF = 1.09, tolerance = 0.92. Participants with prior experience planned the route faster ($M = 80.50$ s, $SD = 65.05$ s) than those without prior experience ($M = 152.20$ s, $SD = 121.63$ s). Nevertheless, the interaction between group (0 = no offloading, 1 = offloading) and additional support use (0 = no support used, 1 = support used) remained statistically significant, β = -1.10, $t(71) = -2.76$, $p = .007$, VIF = 1.62, tolerance = 0.62.

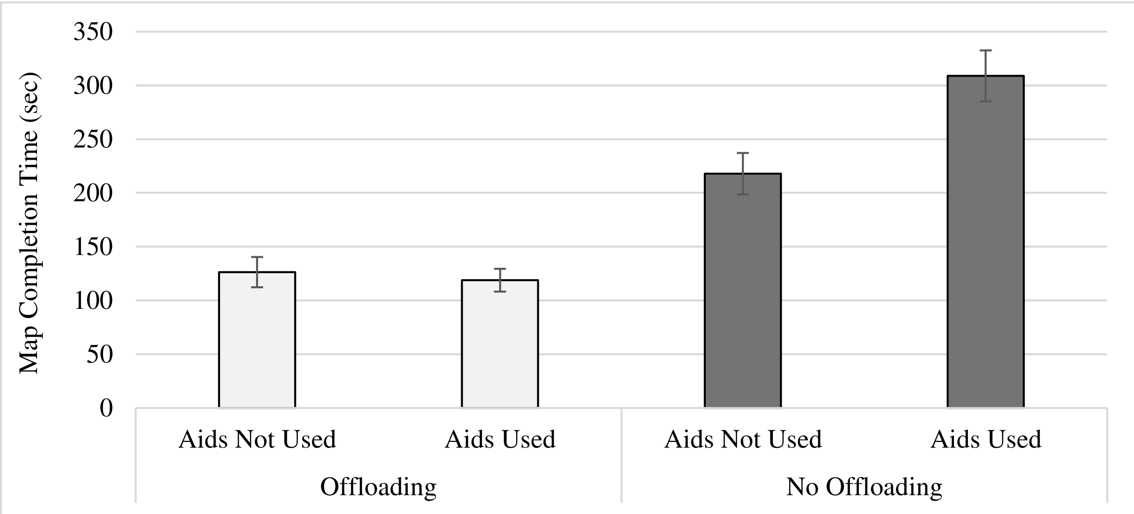
66.38 s, spending significantly more time in typing the final route compared to those who used supports, *M* = 51.71 s, *SD* = 21.90 s.

The performance measures of the route planning task indicate that neither the offloading strategies aided by visual aids embedded into the map, nor the use of additional supports during the experiment, affected planning accuracy. The proportion of participants who completed the map without errors was consistently high across groups (offloading *vs*. no offloading) and the use (*vs*. non-use) of additional supports. Similarly, the presence (*vs*. absence) of offloading strategies supported by visual aids in the maps, or the use (*vs*. non-use) of additional supports during the experiment, did not significantly influence the length of the planned route.

However, both group and additional support use significantly affected time-related measures. Specifically, participants in the offloading group were consistently faster than those in the no offloading group, both in terms of overall time required to complete the map, time devoted to planning the route, and time spent typing the solution. However, in the no offloading group, participants who used additional supports during the experiment spent more time overall to complete the map compared to those who did not use them. The results suggest that this occurred because the use of additional supports, although speeding up solution-typing, slowed down route planning.

Figure 21

Map Completion Time as a Function of Group (Offloading vs. No Offloading) and Additional Support Use (Supports Used vs. Not Used)



Note. Bars represent standard errors.

4.3.3 Route Planning Task Perception

To test the hypotheses that participants in the offloading group would perceive the route planning task as less difficult and less effortful (H2) than those in the no offloading group, we conducted two separate multiple linear regression analyses, one for each perception measure.

Similar to the performance measures, regression models were employed for analysing task perception outcomes because group and additional support use were correlated, and the distribution of participants across the different combinations of these two variables was uneven (Cohen et al., 2003). As before, a hierarchical approach was adopted, with group entered in the first block, additional support use in the second block, and their interaction term in the third block.

Perceived Difficulty of the Route Planning Task

The regression model including only group explained a significant amount of variance in task perceived difficulty, R^2 = .25, F(1,91) = 30.31, p < .001. The inclusion of additional support use significantly increased the explained variance, ΔR^2 = .06, F(1,90) = 7.88, p = .006 (R^2 = .31, F(2,90) =

20.24, $p < .001$). The final addition of the group by additional support use interaction did not further increase the explained variance, $\Delta R^2 = .00$, $F(1,89)=0.51$, $p = .477$.

In the retained model, group (0 = no offloading, 1 = offloading) was a significant predictor of perceived difficulty, β = -1.12, $t(90) = -6.24$, $p < .001$, VIF = 1.07, tolerance = 0.94, indicating that participants in the offloading group ($M = 2.78$, $SD = 1.07$) perceived the test map as less difficult compared to participants in the no offloading group ($M = 4.11$, $SD = 1.24$). Additional support use (0 = supports not used, 1 = supports used) was also a significant predictor, β = -.56, $t(90) = -2.81$, $p = .006$, VIF = 1.07, tolerance = 0.94, with participants who did not use additional supports ($M = 3.55$, $SD = 1.39$) perceiving the task as more difficult compared to those who used supports ($M = 3.22$, $SD = 1.15$).

These findings align with H2. Regardless of the use of additional supports during the experiment, participants who were provided with visual aids embedded in the map to support offloading strategies perceived the task as less difficult than participants who were not provided with such aids. Furthermore, regardless of the presence of visual aids supporting offloading strategies in the map, participants who reported using additional supports during the experiment perceived the task as less difficult compared to those who reported not using such additional supports.

Perceived Effort of the Route Planning Task

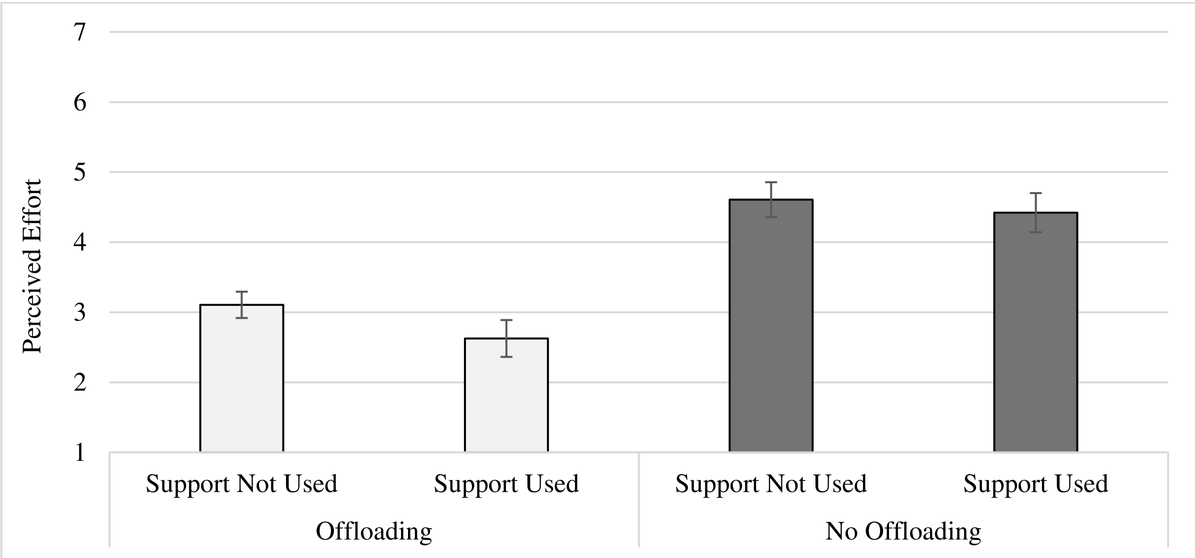
The regression model that included only group as predictor was found to best explain the variance in perceived effort, $R^2 = .29$, $F(1,91) = 37.49$, $p < .001$. Neither the inclusion of additional support, $\Delta R^2 = .01$, $F(1,90) = 1.10$, $p = .297$, nor the addition of the interaction term between group and additional support use, $\Delta R^2 = .00$, $F(1,89) = 0.25$, $p = .616$, further increased the explained variance.

In the best-fitting model, group (0 = no offloading, 1 = offloading) was a significant predictor of perceived effort, β = -1.07, $t(91) = -6.12$, $p < .001$, indicating that participants in the offloading group ($M = 3.02$, $SD = 1.11$) perceived the test map as requiring significantly less effort to be completed compared to participants in the no offloading group ($M = 4.53$, $SD = 1.27$). See Figure 22 for a graphical representation of findings on task perceived effort.

These findings align with H2. Participants who were provided with visual aids embedded in the map perceived the task as less effortful compared to participants who were not provided with such aids. The use of additional supports during the experiment did not lead to a reduction in perceived task effort.

Figure 22

Task Perceived Effort as a Function of Group (Offloading vs. No Offloading) and Additional Supports Use (Supports Used vs. Not Used)



Note. Bars represent standard errors.

4.3.4 Recall of the Planned Route

To test the hypotheses that participants in the offloading group would exhibit better verbal recall of their planned route (H3a), while participants in the no offloading group would demonstrate better visual recall (H3b), we performed a series of multiple linear regression analyses. As in the analyses of performance and perception measures, we used regression models instead of ANOVAs due to the flexibility of regression when estimating effects in the presence of correlated independent variables and uneven participant distribution across the combinations of these variables (Cohen et al., 2003).

A hierarchical approach was again adopted for the recall measures. Group was entered in the first block of each regression, additional support use was added in the second block, and their interaction was included in the third block. Differently from previous analyses, a fourth block was added to exploratively control for the potential influence of the presentation order of visual and verbal recall tasks.

As we did not have specific hypotheses regarding recall time measures, and given that recall time is generally not considered a key variable of interest in such tasks (e.g., Karpicke & Roediger, 2008), the following sections report analyses only for recall accuracy. Analyses of recall time measures are provided in Appendix D.

Verbal recall of the route planned

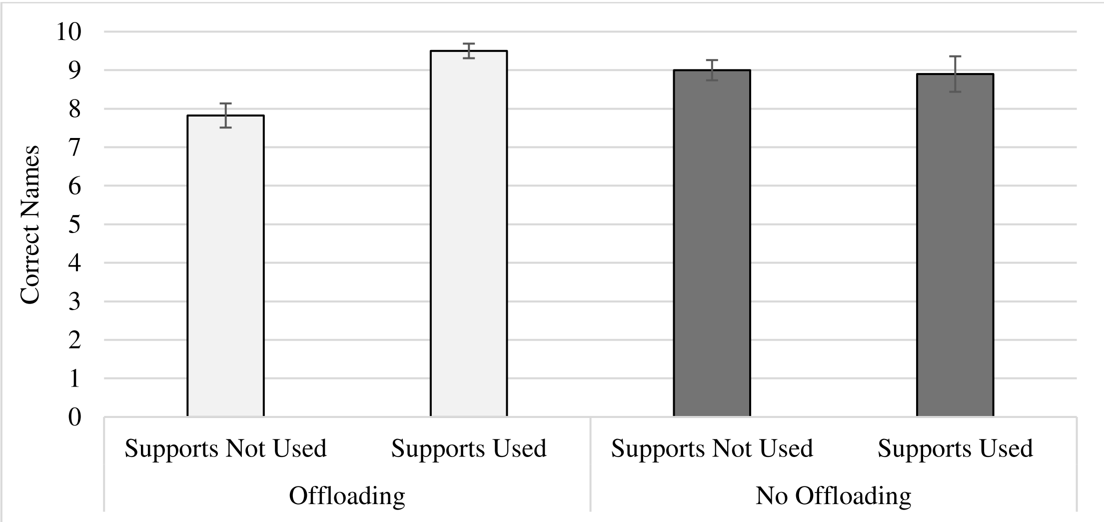
Location Names Correctly Recalled

Concerning the recall of the names associated with the locations in the planned route, the model including only group explained a significant amount of variance, R^2 = .06, F(1,91) = 5.32, p = .023, while the inclusion of additional support use did not significantly increase the explained variance, ΔR^2 = .02, F(1,90) = 1.79, p = .185. However, the inclusion of the interaction between group and additional support use significantly increased the explained variance, ΔR^2 = .04, F(1,89) = 4.42, p = .038 (R^2 = .12, F(3,89) = 3.95, p = .011). The fourth block, controlling for the presentation order of recall tasks, did not increase the explained variance, ΔR^2 = .00, F(1,88) = 0.00, p = .952.

In the third model, the interaction term between group (0 = no offloading, 1 = offloading) and additional support use (0 = supports not used, 1 = supports used) was significant, β = 0.98, t(89) = 2.10, p = .038, VIF = 5.38, tolerance = 0.18, indicating that the effect of group on the recall of location names varied as a function of additional support use (see Figure 23). Further probing of the interaction effect revealed that, among participants who did not use additional supports during the experiment, those in the no offloading group (M = 9.00, SD = 1.39) correctly recalled significantly more location names than those in the offloading group (M = 7.82, SD = 1.93), β = 0.65, t(89) = 2.74, p = .007, VIF = 4.13, tolerance = 0.24. Conversely, among participants who used additional supports, there was no significant difference between the offloading group (M = 9.50, SD = 0.53) and the no offloading group (M = 8.90, SD = 2.01), β= -0.33, t(89) = -0.83, p = .411. Furthermore, among participants in the offloading group, those who used additional supports recalled a greater number of location names compared to those who did not, β =.93, $t(89) = 2.50$, $p = .014$, VIF = 1.69, tolerance = 0.59. Conversely, among participants in the no offloading group, location names recall was not significantly different between those who used and those who did not use additional supports during the experiment, β = -.06, $t(89) = -0.20$, $p = .846$.

Figure 23

Location Names Correctly Recalled as a Function of Group (Offloading vs. No Offloading) and Additional Support Use (Supports used vs. Supports Not Used)



Note. Bars represent standard errors.

Recall of Location Names in the Planned Order

Regarding the location names recalled in their correct order relative to the planned route, the model including only group explained a modest but significant amount of variance, $R^2 = .06$, $F(1,91) = 6.10$, $p = .015$. The inclusion of additional support use significantly increased the explained variance, $\Delta R^2 = .08$, $F(1,90) = 8.73$, $p = .004$. Furthermore, the interaction between group and additional support use also led to a further significant increase in explained variance, $\Delta R^2 = .09$, $F(1,89) = 10.27$, p = .002 ($R^2 = .23$, $F(3,89) = 9.07$, $p < .001$). However, adding the presentation order of recall tasks in the fourth block did not provide an additional contribution to explaining the variance, $\Delta R^2 = .00$, $F(1, 88) = 0.56$, $p = .457$.

In the third model, the interaction between group (0 = no offloading, 1 = offloading) and additional support use (0 = supports not used, 1 = supports used) was significant, β = 1.40, $t(89) = -3.21$, $t(89) = 3.21$, $p = .002$, VIF = 5.39, tolerance = 0.19, indicating that the effect of group on recall accuracy was moderated by the use of additional supports.

Systematic changes in variable reference levels to further probe the interaction effect showed that, among participants who did not use supports, those in the no offloading group ($M = 6.71$, $SD = 2.93$) recalled more location names in the correct order than those in the offloading group ($M = 4.42$, $SD = 2.85$), β = .71, $t(89) = 3.20$, $p = .002$, VIF = 4.13, tolerance = 0.24. However, for participants who used additional supports, there was no significant difference between those in the no offloading group ($M = 7.16$, $SD = 3.32$) and those in the offloading group ($M = 9.38$, $SD = 0.74$), β = -.69, $t(89) = -1.83$, $p = .071$. Notably, among participants in the offloading group, those who used additional supports recalled significantly more location names in the correct order than those who did not ($M = 4.42$, $SD = 2.85$), β = 1.53, $t(89) = 4.43$, $p < .001$, VIF = 1.69, tolerance = 0.59. Conversely, using or not using additional supports did not make a difference in order recall among participants in the no offloading group, β = 0.14, $t(89) = 0.52$, $p = .605$.

Overall, these findings align with H3a. Participants in the no offloading group exhibited a better verbal recall of the planned route, unless participants in the offloading group did use additional external support during the experiment. Notably, the use of such additional supports improved the recall of location names and their order in the planned route only among participants in the offloading group.

Visual recall of the route planned

Location Positions Correctly Recalled

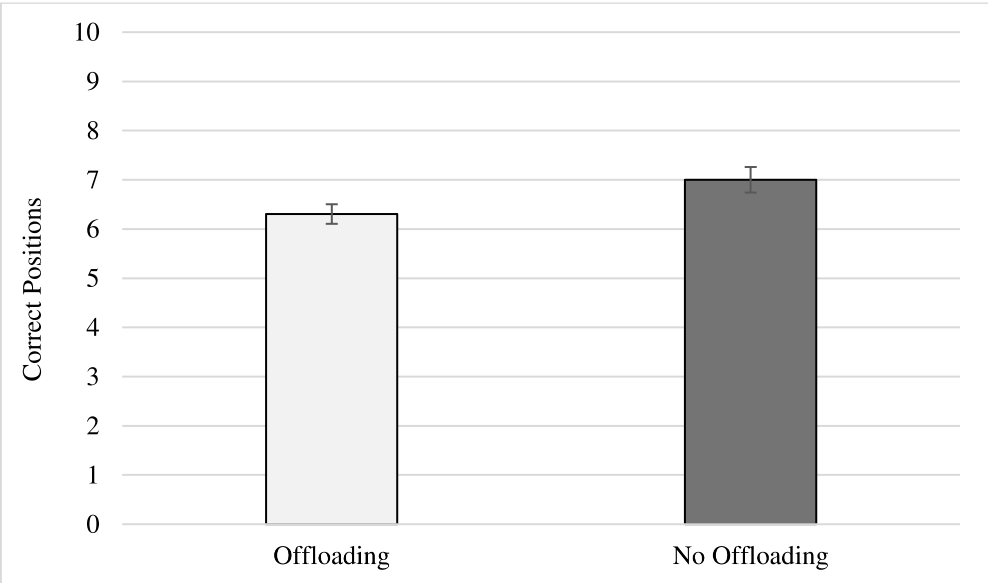
Regarding the location positions included in the planned route correctly recognized (irrespective of their order), the regression model that included only group as a predictor was found to be significant,

$R^2 = .05$, $F(1,91) = 4.50$, $p = .037$. However, the inclusion of additional support use did not significantly increase the explained variance, $\Delta R^2 = .00$, $F(1,90) = 0.05$, $p = .832$, nor did the inclusion of the interaction between group and additional support use, $\Delta R^2 = .00$, $F(1,89) = 0.11$, $p = .738$. Similarly, the inclusion of the presentation order of recall tasks in the fourth block did not contribute to increasing the explained variance of location position recall, $\Delta R^2 = .01$, $F(1,88) = 0.69$, $p = .407$.

In the first model, group (0 = no offloading, 1 = offloading) was a significant predictor of visual recall accuracy, $\beta = -.43$, $t(91) = -2.12$, $p = .037$, indicating that participants in the offloading group ($M = 6.30$, $SD = 1.35$) recognized significantly fewer location positions of their route compared to participants in the no offloading group ($M = 7.00$, $SD = 1.78$). See Figure 24 for a graphical representation of the results on the number location positions correctly recalled by offloading and no offloading groups.

Figure 24

Location Positions Correctly Recalled by Offloading and No Offloading Groups



Note. Bars represent standard errors.

Recall of Location Positions in the Correct Order

The regression analysis on the recall of location positions in their planned order indicated that none of the examined models explained a significant amount of variance. Specifically, the model with only group as a predictor did not significantly account for variance in the recall of location positions in their planned order, R^2 = .01, F(1,91) = 1.22, p = .273. Likewise, adding the use of additional supports did not significantly improve the model, ΔR^2 = .00, F(1,90) = 0.02, p = .900, nor did the interaction between group and additional supports, ΔR^2 =.00, F(1,89) = 0.38, p = .537. The inclusion of the presentation order of recall tasks in the fourth block contributed to a significant increase in explained variance compared to the previous models, ΔR^2 =.05, F(1,88) = 4.27, p = .042, though the overall model remained non-significant, R^2 = .06, F(4,88) = 1.48, p = .215.

In the most parsimonious model, group (0 = no offloading, 1 = offloading) was not a significant predictor of location positions recalled in their planned order, β = -0.23, t(90) = -1.10, p = .273, indicating that participants in the offloading group (M = 5.26, SD = 2.40) did not recall a greater number of location positions in the planned order compared to participants in the no offloading group (M = 5.91, SD = 3.25).

Overall, these findings are inconsistent with the predictions made in H3b. Indeed, participants in the no offloading condition, rather than participants in the offloading condition, showed a better recall of the positions on the map of the locations included in their planned route. However, groups did not differ in their memory of the order in which such locations were included in the planned route.

Paired recall of the route planned

Location Names recalled in the Planned Order

The regression model including only group explained a significant amount of variance in the location names recalled in their correct order during the paired recall task, R^2 = .05, F(1,91) = 4.91, p = .029, and the inclusion of additional support use significantly increased the explained variance, ΔR^2 = .07, F(1,90) = 7.39, p = .008. The interaction between group and additional support use further increased the explained variance, ΔR^2 = .04, F(1,89) = 4.31, p = .041 (R^2 = .16, F(3,89) = 5.81, p = .001) . However, adding the presentation order of recall tasks in the fourth block did not significantly improve the model, ΔR^2 = .00, F(1,88) = 0.05, p = .817.

In the third model, the interaction between group (0 = no offloading, 1 = offloading) and additional support use (0 = supports not used, 1 = supports used) was significant, β =.94, t(89) = 2.08, p = .041, indicating that the effect of group on location names recalled in their correct order during the paired recall task varied as a function of additional support use. Further probing of the interaction indicated that, among participants who did not use additional supports, those in the no offloading group (M = 6.36, SD = 3.37) recalled significantly more location names in the correct order than those in the offloading group (M = 4.42, SD = 3.05), β = .55, t(89) = 2.40, p = .019, VIF = 4.13, tolerance = 0.24. However, among participants who reported additional support use, there was no significant difference between the no offloading (M = 7.26, SD = 3.62) and offloading groups (M = 8.63, SD = 2.72), β = -.39, t(89) = -1.00, p = .322. Remarkably, among participants in the offloading group, those who used additional supports recalled significantly more location names in the correct order than those who did not use these supports, β = 1.20, t(89) = 3.33, p = .001, VIF = 1.69, tolerance = 0.59. Conversely, for participants in the no offloading group, the use of additional supports did not significantly affect recall compared to those who did not use supports, β = .26, t(89) = 0.94, p = .350.

These findings suggest that, in the absence of additional support use, participants in the no offloading group were better able to recall the planned order of location names compared to those in the offloading group, providing further support to H3a. However, this advantage was lost when additional supports were employed, as participants in the no offloading group did not outperform those in the offloading group under these circumstances.

Location Positions recalled in the Planned Order

The regression analysis on the recall of location positions in their correct order during the paired recall task indicated that none of the examined models explained a significant amount of variance in this measure. The model with only group as a predictor did not significantly account for variance in the number of location positions recalled in their planned order, $R^2 = .00$, $F(1, 91) = 0.29$, $p = .594$. Including the use of additional supports did not improve the model, $\Delta R^2 = .00$, $F(1, 90) = 0.01$, $p = .943$, nor did the interaction between group and use of additional supports, $\Delta R^2 = .01$, $F(1, 89) = 0.51$, $p = .475$. Finally, adding the presentation order of recall tasks in the fourth block did not significantly increase the explained variance, $\Delta R^2 = .02$, $F(1, 88) = 2.01$, $p = .159$.

In the most parsimonious model, group (0 = no offloading, 1 = offloading) was not a significant predictor, $\beta = .11$, $t(91) = 0.54$, $p = .594$, indicating that participants in the offloading group ($M = 7.28$, $SD = 2.98$) did not recall significantly more location positions in the planned order during the paired recall task compared to participants in the no offloading group ($M = 6.94$, $SD = 3.25$). These findings are consistent with the results from the visual recall task, but both are inconsistent with H3b. Indeed, contrary to our expectations, participants in the offloading group did not remember the order in which they visited the locations on the map in their route better than participants in the no offloading condition, neither in the visual recall task nor in the paired recall task.

Pairing of Location Names-Positions in the Planned Route

The regression model including only group explained a significant amount of variance in the correct pairings between location names and their positions in the map of the places in the planned route, $R^2 = .06$, $F(1,91) = 5.70$, $p = .019$, while the inclusion of additional support use did not significantly increase the explained variance, $\Delta R^2 = .03$, $F(1,90) = 2.82$, $p = .097$. However, the group by additional support use interaction significantly increased the explained variance, $\Delta R^2 = .05$, $F(1, 89) = 4.67$, $p = .033$ ($R^2 = .06$, $F(1,91) = 5.70$, $p = .019$). The addition of the presentation order of recall tasks did not significantly increase the explained variance, $\Delta R^2 = .01$, $F(1,88) = 0.53$, $p = .470$.

In the third model, the interaction between group (0 = no offloading, 1 = offloading) and additional support use (0 = supports not used, 1 = supports used) was significant, $\beta = 1.00$, $t(89) = 2.16$, $p = .033$, VIF = 5.39, tolerance = 0.19, indicating that the effect of group on correct pairings was moderated by additional support use.

Further probing of the interaction revealed that, among participants who did not use additional supports, those in the no offloading group, $M = 6.43$, $SD = 3.16$, made significantly more correct pairings than those in the offloading group, $M = 4.26$, $SD = 3.05$, $\beta = .65$, $t(89) = 2.77$, $p = .007$, VIF = 4.13, tolerance = 0.24. However, among participants who used additional supports, there was no significant difference between the no offloading, $M = 6.47$, $SD = 3.50$, and offloading groups, $M = 7.63$, $SD = 2.45$, $\beta = -.35$, $t(89) = -0.87$, $p = .386$. Interestingly, among participants in the offloading group, those who used additional supports made significantly more correct pairings than those who did not use them, $\beta = 1.02$, $t(89) = 2.76$, $p = .007$, VIF = 1.69, tolerance = 0.59. Conversely, for participants in the no offloading group, the use of additional supports did not significantly affect pairing performance compared to those who did not use them, $\beta = .01$, $t(89) = 0.05$, $p = .961$.

These findings suggest that participants in the no offloading group made more accurate name-position pairings of the locations included in the planned route compared to participants in the offloading group. However, this was observed only when no additional supports were used during the experiment. Indeed, the use of additional supports improved performance in the offloading group, eliminating the gap between groups.

Name-Position Pairings in the Planned Order

Concerning the correct pairings recalled in their planned order during, the regression analysis revealed that the model including only group explained a significant amount of its variance, $R^2 = .05$, $F(1,91) = 4.29$, $p = .041$. The inclusion of additional support use did not significantly improve the explained variance, $\Delta R^2 = .03$, $F(1,90) = 2.57$, $p = .112$, nor did the addition of the group by additional support use interaction, $\Delta R^2 = .02$, $F(1,89) = 1.94$, $p = .167$. Finally, the inclusion of the presentation order of recall tasks in the fourth block also did not significantly improve the model, $\Delta R^2 = .01$, $F(1,88) = 1.34$, $p = .249$.

In the first model, group (0 = no offloading, 1 = offloading) was a significant predictor of correct name-position pairings recalled in the planned order, $\beta = -.42$, $t(91) = -2.07$, $p = .041$, indicating that participants in the offloading group, *M* = 4.41, *SD* = 3.14, recalled significantly fewer correct pairings in their planned order than those in the no offloading group, *M* = 5.85, *SD* = 3.54.

Therefore, regardless of additional support use, participants in the no offloading (*vs*. offloading) group exhibited better memory for the names associated with the locations in the map and, specifically, they remembered better the order in which they had included the locations and their associated names in their route.

4.4 Discussion

Study 4 aimed to investigate whether the offloading strategies spontaneously devised for route planning in Study 1 and Study 2, and directly supported by visual aids embedded into the maps in Study 3, would remain effective when applied to more ecologically valid materials compared to the maps previously used. For this purpose, we employed a map of a real-world location.

Overall, findings from Study 4 indicate that providing visual aids for offloading strategies supported participants' route planning even in this more realistic context, partly in line with H1. Specifically, participants provided with aids for cognitive offloading were much more efficient in planning their route compared to those without such aids, without any loss in accuracy. Additionally, the presence of these aids into the map reduced perceived task difficulty and effort, supporting our second hypothesis (H2). However, the findings for route recall were only partially in line with our expectations. Indeed, participants in the no offloading group generally demonstrated better recall of the names in the planned route, as predicted by H3a, but also showed better recall of the letters indicating the positions in the maps included in their planned route, which contrasts with H3b.

The proportion of participants who completed the real-world map without errors was consistently high and comparable between the offloading and no offloading groups. Furthermore, the presence of external visual aids did not enhance route planning accuracy, as participants in the offloading group did not deviate less from the shortest route compared to those in the no offloading group. Although these findings do not support H1, they are generally consistent with the results of our previous studies.

In Studies 1 to 3, we also found that the opportunity to offload cognition during route planning did not significantly impact error commission rates. This may be because participants were generally able to avoid mistakes even without the option to offload cognition. Concerning the goodness of the planned routes, only in Study 1 we did observe that participants provided with the opportunity to offload cognition deviated less from the optimal (shortest) ones compared to participants without such opportunity. In Studies 2 and 3, the removal of the possibility to offload cognition led to a greater length of planned routes, while providing this possibility resulted in slight improvements only. However, during the training phase of both the studies, no significant differences in route length were observed between participants who could offload cognition and those who could not. A possible explanation for these mixed findings is that cognitive offloading may only improve route planning accuracy when participants are required to plan a relatively large number of routes, and the reliability of the task is possibly increased. In Study 1, participants planned routes on a total of 16 maps, whereas Studies 2 and 3 progressively reduced the number of maps, with Study 4 using only a single map. This pattern suggests that the benefits of cognitive offloading for route planning accuracy may become more apparent when the task involves planning multiple routes across different maps over a relatively extended time span, or when participants shift from the possibility of offloading cognition (and thus relying on a visual strategy for planning) to its removal, which requires an abrupt transition to a more demanding strategy, primarily based on working memory.

However, participants with visual aids for offloading strategies on the real-world map completed the task in approximately half the time compared to those without visual aids, supporting H1 and extending previous findings to more ecologically valid materials. A more in-depth analysis of time-related measures of route planning indicated that participants in the offloading group were faster than those in the no offloading group both in planning the route and in reporting it once fully identified. Interestingly, within the offloading group, participants who reported using additional supports did not complete the map significantly faster or slower compared to those who did not use such supports. However, in the no offloading group, participants who used additional supports spent more time completing the map compared to those who did not, expanding their disadvantage relative to participants in the offloading group. This increased time was primarily due to spending more time planning the route, although once the route was identified, they became faster at reporting it.

We checked for the use of additional supports such as paper and pencil, computers, or smartphones, as their use indicates that participants may have engaged in some form of cognitive offloading during the experiment beyond the strategies supported by the visual aids. The fact that participants in the offloading condition did not further improve their map completion time when using additional supports suggests they engaged in a more than optimal use of cognitive offloading, aligning with results from our earlier studies. Specifically, Studies 1, 2 and 3 extended previous findings on cognitive offloading (Gilbert et al., 2020; Gilbert et al., 2023), showing that over-reliance on offloading also occurs in route planning. Study 4 further extends these findings by indicating that individuals may exhibit a similar bias toward cognitive offloading even when planning a route on a real-world map. The fact that participants in the no offloading group experienced further slowdown when using external supports provides additional evidence for this idea, while also suggesting the greater efficiency of the visual aids for offloading strategies we embedded into the maps compared to other external aids for cognitive offloading.

Regardless the use of additional supports during the experiment, participants who planned the route with visual aids for offloading strategies perceived the task as less difficult and effortful (H2), consistently with our second hypothesis. Notably, even beyond the presence of visual aids in the map, the use of additional supports during the experiment was associated with lower levels of perceived task difficulty compared to when no additional supports were used. Given that employing additional supports did not improve route planning performance among participants in the offloading group, and even slowed it down among participants in the no offloading group, this finding suggests that individuals may rely on forms of cognitive offloading not necessarily to enhance performance, but rather to reduce cognitive effort and/or alleviate the subjective experience of task difficulty. This would align with findings from Studies 2 and 3, which indicated that greater reliance on cognitive offloading, even when it did not improve task performance, still provided benefits in terms of reducing effort or perceived task difficulty. Thus, while cognitive offloading may not always lead to objective performance gains, its impact on task perception may play a key role in whether individuals choose to rely on internal or external resources when completing a complex task.

When no additional supports were used, participants in the no offloading group generally exhibited better verbal recall of the planned route compared to participants in the offloading condition. Specifically, in the task requiring verbal recall of the planned route, participants in the no offloading group more accurately recalled both the names of the locations they included in their route and the order in which they planned to visit them. Similarly, in the paired recall task (which required recalling the planned route both verbally and visually), participants in the no offloading group were still better able to recall the planned order of location names compared to those in the offloading group. These findings are consistent with H3a and further support the idea that, in the absence of external aids for cognitive offloading, participants rely on an incremental strategy, primarily based on working memory, to plan their route, as suggested by our previous studies. Specifically, we expected participants in the no offloading group to progressively add the locations to be visited to their planned route while mentally rehearsing their names in sequence, leading to a deeper encoding of both the names and their order. The better verbal recall of the planned route by participants in the no offloading condition provide further evidence to this interpretation.

However, the advantage in verbal recall of the planned route shown by participants in the no offloading group disappeared in both the verbal recall and paired recall tasks when additional supports were employed during the experiment. Under these circumstances, participants in the no offloading group no longer exhibited better recall of location names or their positional order within the planned route compared to participants in the offloading group. Although we cannot determine exactly how participants used external supports during the experimental tasks, these findings suggest that participants who reported using additional supports may have kept an external record of their route while planning it, possibly by writing down some of the location names. This behaviour would mirror the option provided in the full offloading condition of Study 3, where a textbox was available for annotations during route planning. Although speculative, this inference arises from the consideration that, if the use of additional supports was restricted to the recall tasks, participants should not have benefitted from them during verbal recall of the planned route. Indeed, participants were unaware that they would need to recall the route later, and using external supports only during the recall phases would likely have been ineffective. Additionally, the fact that the use of external supports affected recall of the names of places in the planned route, but not the recall of the positions of the locations included in the planned path, supports the idea that participants may have used these supports specifically to track the location names. Notably, only participants in the offloading condition benefitted from the use of additional supports when recalling the planned route verbally, while participants in the no offloading condition did not gain further advantage from using them. This suggests that the initial encoding of the names of the locations in the planned route by participants in the no offloading group was sufficiently deep to support recall without the need for additional external aids.

Contrary to H3b, participants in the no offloading condition, rather than those in the offloading condition, demonstrated better recall of the positions on the map of the locations that were included in the route. However, the two groups did not differ in their memory of the order in which these locations were included in the planned route, neither in the task requiring the recall of the positions in the route nor in the task requiring both name and position recall. We hypothesized that participants in the offloading group would exhibit better recall of the positions in the planned route, as we expected them to rely on a predominantly visual strategy to support their planning. Specifically, given that these participants could offload the need to remember locations and ordering rules onto the visual aids, we anticipated that they would focus more on the visual aspects of the map, leading them to better remember the visuospatial characteristics of the planned route.

One possible explanation for this unexpected finding is that, contrary to our expectations, participants in the offloading group may have engaged in a superficial encoding of the visuospatial aspects of the map, due to the visual aids highlighting the target locations and ordering rules. By reducing the noise from distractor locations, these visual aids might have made route planning more perceptually guided, decreasing the need for extended encoding of specific visuospatial characteristics of the map, like the precise positions of the locations in the planned path. As a result, while the route planning could be efficiently supported by the visual aids, participants in the offloading group may not have deeply encoded the specific positions of the target locations on the map.

An alternative, though not mutually exclusive, explanation is that participants in the no offloading group, lacking visual aids to support their planning, needed to mentally maintain the positions of the locations on the map during planning, while also inhibiting interference from distractor locations. This may have resulted in a better encoding of the spatial positions of the locations to be visited, even if it did not lead to a better memory of their order of visit.

These interpretations do not necessarily contradict the idea that participants in the offloading group used a predominantly visual strategy for planning. However, they suggest that the visual aids may have altered the nature of the route planning task, leading to a more surface-level encoding of the map. Research on traveling salesman problems (TSPs) supports this notion, as it indicates that individuals tend to employ an effective global perceptual strategy in TSPs, which closely resembles route planning under offloading conditions (Kong & Schunn, 2007; Ormerod & Chronicle, 1999; van Rooij et al., 2006). When visual aids are embedded in the map, distractors do not interfere with route planning, just as in route planning in TSPs, where no distractors are provided. This similarity suggests that the presence of visual aids may have promoted a more global perceptual strategy, reducing the need for deeper encoding of specific visuospatial features. Further research is needed to clarify the impact of different offloading strategies on the depth of encoding of spatial positions in our task.

Although the findings on recall measures only partially align with our third set of hypotheses, they are consistent with previous research showing the negative effects of cognitive offloading on episodic memory (e.g., Grinschgl et al., 2021). Indeed, the available evidence suggests that relying on external aids can impair memory for offloaded information, whether verbal (e.g., Sparrow et al., 2011) or visual (e.g., Henkel, 2014). Some authors argue that this occurs because, once information is offloaded onto external aids, individuals may shift their focus away from encoding the content itself (Henkel et al., 2014; Sparrow et al., 2011). Similarly, research on spatial memory has shown that navigational aids can weaken memory for environmental details and navigated routes (Fenech et al., 2010; Gardony et al., 2015). These parallels suggest that, while external aids for offloading strategies are effective in improving route planning performance and reducing effort and perceived task difficulty, they may also impair the encoding of information, which aligns with the patterns observed in our study.

Further support for this account comes from the findings of the paired recall tasks. When no additional supports were used, participants in the no offloading group correctly paired a greater number of location names to their positions on the map compared to participants in the offloading group. However, when additional supports were used, this gap disappeared, but only participants in the offloading group benefitted significantly from them. Remarkably, participants in the no offloading group better recalled the planned order of correctly paired name-position locations than those in the offloading group, regardless of the use of additional supports.

Despite the valuable insights provided by Study 4, several limitations should be considered when interpreting the findings. More than half of the participants had previously participated in a study involving route planning on maps. However, participants were evenly distributed between the two groups and their previous experience did not affect the main findings related to the comparison between the offloading and no offloading group.

Nearly one-third of participants reported using additional supports during the experiment, with a higher frequency in the no offloading group. While we controlled for this factor in our analyses, it is possible that this reduced the statistical power of our results. This finding, however, further supports the idea that participants tend to spontaneously employ external aids when faced with complex tasks, such as map-based route planning. While this did not lead to performance improvements, it did reduce participants' perception of task difficulty.

The question about the use of additional supports was ambiguous, as it did not specify at which stage of the experiment participants used external aids or which specific aids have been used. However, our findings seem to suggest that participants may have kept a track of the route while planning, likely by writing down location names. This interpretation is supported by the fact that participants in the no offloading group took significantly longer to complete the map when they used external aids and that participants in the offloading group improved their recall of the names in the planned route when using additional supports, while this did not extend to the recall of the letters indicating their positions.

The finding that participants in the present study tended to spontaneously adopt offloading strategies during complex tasks may present a limitation also for Study 3, as it was conducted online, but without controlling for the potential use of additional supports. Future online experiments on cognitive offloading should systematically control for the potential influence this factor.

We did not include the option to take annotations during route planning in the offloading condition of Study 4. This decision was motivated by the inconsistent use of this offloading strategy in Study 3 and the fact that it did not further improve route planning performance when employed. However, our findings indicate that participants in the offloading condition still used additional supports, and they may have specifically taken notes on the location names in the planned route, as suggested by their better recall of these names when using external aids. If our interpretation of the use of additional supports is correct, it aligns with the findings from Studies 1 and 2, which indicated that tracking the route during planning is a naturally adopted offloading strategy. Nonetheless, as seen in Study 3, this strategy does not significantly enhance performance.

Overall, Study 4 shows that cognitive offloading strategies improve efficiency in route planning when applied to an ecologically valid material, such as a real-world map. Visual aids embedded in the map reduced task completion time and perceived effort and difficulty , consistently with our findings from previous studies, and did not reduce accuracy. However, the results also suggest that reliance on offloading strategies may come at a cost to memory performance. These findings contribute to the growing body of literature on cognitive offloading, providing further evidence of its benefits when facing with complex and realistic tasks, while also highlighting the potential cognitive trade-offs involved in relying on external supports. In terms of applied implications, these results suggest that the visual aids we used to support cognitive offloading could improve performance in real-world tasks such route planning. It follows from this that they may have the potential to inform the design of improved navigational systems or assist individuals with deficits in route planning. However, careful consideration is needed in contexts where memory retention of the offloaded information is critical. Balancing the use of external aids with the need for internal cognitive resources may be key to designing tools that enhance efficiency without undermining important cognitive processes.

GENERAL DISCUSSION

The studies presented in this thesis aimed to expand research on cognitive offloading by investigating it in the previously unexplored domain of route planning. They also sought to increase current knowledge of the consequences of cognitive offloading, particularly by investigating the transfer effects of its use or non-use on subsequent route planning performance.

One of the key contributions of the studies presented in this thesis lies in the observation that participants simultaneously used a variety of strategies to offload cognition in our complex task. The integrated use of multiple offloading strategies has not been explored in prior research, which primarily studied the effects of using a single offloading strategy at a time, such as setting reminders in intention offloading tasks (e.g., Gilbert et al., 2023), opening and closing windows in pattern copy tasks (e.g., Grinschgl et al., 2020), or offloading information onto digital devices (e.g., Henkel et al., 2014; Sparrow et al., 2011). However, our findings suggest that, in complex tasks with high cognitive demands, people may naturally rely on a combination of strategies to manage cognitive load. The integrated use of multiple offloading strategies warrants further investigation. It would be valuable to understand whether it can be observed in other complex cognitive tasks beyond route planning, as this could deepen our understanding of how offloading strategies are employed in real-world contexts. Future studies could also explore the mechanisms underlying the selection and integration of offloading strategies, as well as the conditions under which an integrated use provides greater benefits to performance than using a single strategy.

Our studies also extend some of the main results observed in previous research on cognitive offloading to the route planning domain, while challenging others. Consistent with earlier studies, we found that cognitive offloading was used more frequently when the task imposed a greater cognitive load (e.g., Gilbert, 2015a; see Study 1) and we observed that its use improved performance (e.g., Risko & Gilbert, 2016; see Studies 1, 2, 3, and 4), but only when the task was more difficult (e.g., Risko & Dunn, 2015; see Study 1). We also observed that participants tended to continue using offloading strategies after prior use (e.g., Scarampi & Gilbert, 2020; see Studies 1 and 2) and to rely on them more than needed for performance benefits (see Studies 1, 2; see also Study 3). This is consistent with previous studies showing that individuals tend to rely on cognitive offloading in a suboptimal way (e.g., Gilbert et al., 2020). As we observed that some of the offloading strategies were used without further performance gains, this could also inform metacognitive training aimed at encouraging their more adaptive and effective use (e.g., Engeler & Gilbert, 2020).

However, Studies 2 and 3 also indicated that this overreliance on cognitive offloading, while not providing additional performance advantages, still offered some benefits in terms of perceived effort and task difficulty . These findings align with previous accounts emphasizing the role of the avoidance of cognitive effort in triggering cognitive offloading (e.g., Sachdeva & Gilbert, 2020) and, more generally, they suggest that the bias toward cognitive offloading over internal resources can also be driven by participants' need to reduce perceived task effort or difficulty, rather than solely by the need to support performance.

However, we also found results contrasting with the previous literature. Specifically, while prior evidence indicates that lower metacognitive confidence in task performance predicts a greater reliance on cognitive offloading (Gilbert et al., 2020; see also Gilbert et al., 2023; Grinschgl et al., 2020; Risko & Dunn, 2015), we did not observe a significant relationship between participants' performance expectancies in our route planning task and their use of offloading strategies (see Study 1). Moreover, while previous studies found that lower cognitive abilities were related with a greater use of cognitive offloading (e.g., Gilbert, 2015a; Meyerhoff et al., 2021), we observed the opposite: better, rather than worse, cognitive abilities were associated with a greater use of cognitive offloading. Specifically, better visual search abilities were correlated with greater use of offloading strategies in Study 1 and Study 2. Study 3 corroborated these findings by showing that better visual search abilities were associated with shorter routes when visual aids for all offloading strategies were embedded in the maps. Our results suggest that the offloading strategies that participants spontaneously devised in Studies 1 and 2 (externally supported in Studies 3 and 4) may have been used to tailor the task to better fit with participants' cognitive strengths. Notably, when no offloading options were provided, we found that route planning performance was correlated with spatial working memory in Studies 1 and 2, and with verbal working memory in Study 3. Therefore, our studies suggest that, by using offloading strategies, participants did not simply reduce task cognitive demands. Rather, they appear to have changed the nature of the task itself, transforming it into a primarily visual one. As a result, the cognitive abilities underlying route planning performance shifted from working memory to visual-perceptual skills. The discrepancy between our studies and previous literature on the role of individual differences in cognitive abilities highlights the need for further research into how such differences influence the use of offloading strategies in complex tasks, as this relationship appears more complex than previously hypothesized (see also Meyerhoff et al., 2021). Specifically, further studies are needed to clarify under which conditions better cognitive abilities are associated with greater cognitive offloading, and when the reverse happens.

It is also interesting to note that, in Study 3, route planning performance was correlated with visual search abilities only in the condition in which visual aids for all offloading strategies were provided, whereas no such correlation was observed when aids were provided for only a subset of them. This result might suggest that participants in the partial offloading condition employed a hybrid strategy for planning routes, rather than relying on a predominantly visual strategy. While this interpretation remains speculative, the finding further highlights the need to explore the integrated use of multiple offloading strategies more in depth.

To our knowledge, Studies 2 and 3 are the first to simultaneously explore bidirectional transfer effects in the field of cognitive offloading. Their results offer new insights into the consequences of cognitive offloading, while also expanding upon the existing literature on transfer effects. We hypothesized that shifting from a primarily visual strategy (with offloading) to a strategy primarily based on working memory (without offloading) would be more challenging than the reverse. Consistently with this, both studies showed that removing cognitive offloading after training with it led to a greater performance decline than the improvement seen when offloading was introduced after training without it. These findings extend prior evidence on the negative side-effects of cognitive offloading (e.g., Sparrow, 2011; Henkel, 2014; Fenech, 2010), further demonstrating the detrimental impact of prior use of offloading in a route planning task on subsequent unaided task performance. Importantly, both studies also showed that participants quickly adapted to cognitive offloading after having experienced the task without using it.

These results support a refined view of strategy shifts within a task than the one provided by previous studies on transfer effects in problem-solving (Thomas et al., 2018; Bilalić et al., 2008, 2008b, 2010; Luchins & Luchins, 1950) and by the studies on the positive consequences of training with a planful strategy prompted by high operator cost (O'Hara & Payne, 1998). Indeed, Studies 2 and 3 indicate that the ease or difficulty of the shift between planning strategies is contingent upon the cognitive features of both the initial and final strategies. More broadly, our results point out to the importance of considering the specific interaction between the strategies used by participants over time. Further research should examine whether these asymmetric transfer effects hold across different types of cognitive tasks and offloading strategies.

Studies 2 and 3 also revealed that individual differences in working memory moderated asymmetric transfer effects. In Study 2, participants trained with pen use better dealt with its removal when they had better spatial working memory skills, as they experienced a less pronounced performance decline. Similarly, in Study 3, participants trained with visual aids for offloading strategies showed a smaller slowdown in solution-typing time after the removal of these aids when they had better verbal working memory skills. Beyond reinforcing the idea that working memory supports route planning when cognitive offloading is unavailable, these results also suggest that working memory eased the transition to strategies requiring internal cognitive resources of participants initially trained with the opportunity to offload cognition. Thus, individual differences in working memory seem to ease the shift between externally-aided and internally-driven strategies, an overlooked aspect that warrants further investigation.

Study 4 extended some of the findings of previous studies, which employed artificial maps, by using a map of a real place. Specifically, Study 4 demonstrates that providing visual aids embedded in a real-world map for the offloading strategies we identified was effective in supporting route planning performance. This is a useful contribution, as most of previous studies on cognitive offloading used rather simple laboratory tasks (e.g., Intention Offloading task, Gilbert, 2015a; Pattern Copy task, Grinschgl et al., 2020). Study 4 aligns with the emerging trend in cognitive offloading research towards the investigation of this phenomenon in more realistic contexts. For example, Gilbert et al. (2024) examined cognitive offloading with ChatGPT in a study in which participants could opt for AI assistance over self-generated responses to quizzes and creativity tasks. Surprisingly, they found that participants tended to answer by themselves, even though using ChatGPT would have yielded better outcomes. This tendency was driven by participants' overconfidence in their abilities, with further evidence of a general bias either for or against ChatGPT shaping their reliance on the AI. McDonald and Soares (2024) observed a "digital fixation" effect with AI tools, specifically finding that participants who had previously used Google or ChatGPT to answer difficult trivia questions continued to rely on these tools for easier questions, despite having the option to use their own memory. Remarkably, the AI tool chosen under free choice consistently matched the one previously used. This work thus extends the perseverative and more-than-optimal use of prior offloading strategies to behaviors in realistic digital contexts. Reflecting this trend, Grinschgl and Venus (2024) have recently developed a questionnaire to assess everyday offloading behaviors aiming to provide a foundation for studying individual differences in offloading habits in real-world settings and their predictors. Study 4 contributes to this more ecologic line of offloading research by providing evidence supporting the efficiency of visual aids for offloading strategies in route planning on a real-world map.

A further contribution of Study 4 lies in its corroboration of previous evidence on the negative impact of cognitive offloading on episodic memory (Grinschgl et al., 2021, Henkel et al., 2014; Sparrow et al., 2011). Indeed, we found that participants without offloading aids on the maps demonstrated better recall of the planned route compared to those with such aids. Notably, these findings partly contradict one of our hypotheses. Indeed, our expectation that participants without offloading aids would demonstrate better recall of location names, as using a working memory-based strategy was anticipated to enhance their encoding through internal rehearsal, aligned with our findings. However, our expectation that participants provided with offloading aids would show better recall of the position of the locations in the planned route, given that a strategy mainly based on visual planning was expected to improve the encoding of positions on the map, was inconsistent with the results. Nonetheless, these findings further reinforce the notion that the availability of offloading aids changed the nature of our task. In particular, participants who were provided with aids for offloading strategies during route planning may have transformed our task in a sort of traveling salesman problem (TSP). TSPs requires to find the shortest path among a set of points in the absence of distractors and rules, and individuals typically solve them very effectively by adopting intuitive global visual-perceptual strategies (Kong & Schunn, 2007; van Rooij et al., 2006; Ormerod & Chronicle, 1999). The availability of offloading aids may have made the to-be-visited locations and rules to follow visually salient over distractors, prompting such a global visual-perceptual approach. This may have eventually resulted in a more superficial encoding of the specific spatial positions of the locations included in the route during planning.

Overall, the studies presented in this thesis may represent a promising approach to be followed in the investigation of cognitive offloading for theoretical and applied aims. This approach includes the identification of the offloading strategies spontaneously devised in response to complex tasks (Study 1), the assessment of their relationships with performance (Study 1 and Study 2), the evaluation of the supporting role of visual aids for offloading strategies embedded in the task (Study 3), and the test of these aids in a real-world context (Study 4). A similar approach could be extended to other tasks and cognitive domains.

Moreover, our approach has also the potential to guide the development of assistive devices or programs for people with or at risk of cognitive vulnerabilities, as recent evidence in the literature indicates that offloading aids may be beneficial for such populations. Insightful in this regard is the recent work by Scullin et al. (2024), who found that individuals with mild cognitive impairment successfully learned to offload intentions using a smartphone reminder app across a 4-week intervention period, thereby compensating for their prospective memory deficits. In line with this, a recent meta-analysis by Benge and Scullin (2023) indicated that, among older adults, engagement with everyday digital technologies is associated with decreased risk of cognitive decline and dementia, suggesting that using technological offloading aids in daily life may contribute to building a "technological reserve". In line with these results, the offloading strategies we identified and the visual aids we tested could be incorporated into training or rehabilitation programs for route planning or spatial planning abilities (e.g., Basso et al., 2001).

Finally, the offloading strategies we identified and the visual aids we tested may help to enhance currently available navigation systems. Free navigation tools like Google Maps, while highly effective for finding paths between two points, typically lack functionality for optimizing the sequence of multiple destinations or incorporating specific constraints between locations (e.g., time restrictions, preferred visit order, or restricted routes between points of interest). Changing these systems and their interfaces to match the way in which participants naturally offload route planning on maps and capitalize on the respective strengths of humans and systems may improve system functionality and usability.

Certain limitations of our studies must be addressed to fully evaluate the scope and implications of our findings. In Study 2, we investigated transfer effects by either removing or introducing the possibility to offload cognition following training with or without such a possibility. However, the extent of decline/improvement in route planning performance after removal/introduction of cognitive offloading was assessed without including control groups that consistently completed the training and test phases of the task with or without cognitive offloading. Study 3 partially addressed this limitation by including a control group that completed both the training and test phases without offloading. This addition allowed us to further corroborate the notion that prior offloading experience negatively affected subsequent task performance without offloading, as we observed that, in the test block without offloading aids, the groups trained with them showed worse route planning performance compared to the control group. However, Study 3 did not include a control group that performed the task with visual aids for offloading strategies available from the training to the test phases. Therefore, we cannot determine whether participants trained without offloading aids, once provided with them, would have performed comparably, better, or worse compared to this missing control group. Future studies should include both control groups to gain a more complete understanding of transfer effects within the context of cognitive offloading.

Notably, approximately one-third of participants in Study 4 reported using external supports, such as paper and pencil, smartphones, tablets, or computers, to help themselves to complete the experiment, particularly in the group without visual aids embedded in the maps. This aspect was not controlled for in Study 3, which was also conducted online. Online studies of cognitive offloading are limited in that they do not allow control on the use of offloading strategies beyond those designed experimentally, or control on the use of offloading strategies in experimental conditions designed to preclude them. To our knowledge, Study 4 is the first online study on cognitive offloading to formally control for this aspect, and its findings highlight the importance of including this control in future online studies.

Another limitation is that participants to our studies were undergraduate students. Future studies could usefully expand our investigation to other age groups, with particular attention to older populations. Evidence suggests that older adults might underutilize offloading strategies due to an overestimation of their unaided abilities (Scarampi & Gilbert, 2021). Yet, it is in this age group that offloading aids may be particularly beneficial, as they can offer key supports for daily functioning and autonomy (e.g., Scullin et al., 2020; Benge & Scullin, 2023).

It should also be considered that, nowadays, planning a route on a physical map may be a rather infrequent behaviour, as we all have quick access to digitised maps via a smartphone, tablet or computer. Studies 3 and 4 attempted to account for this by moving away from a paper-and-pencil approach to presenting digital maps and visual aids for offloading. However, future studies should take additional steps to investigate the offloading strategies that individuals spontaneously adopt in interactive digital contexts that more closely reflect current real-world map use, such as route planning aided by navigation technologies. It would also be beneficial to observe whether some offloading behaviours, although spontaneously devised on paper maps, are not implemented because they are constrained by software design, as this would help to improve such technologies from a user-centred perspective.

Finally, while our studies provided valuable insights into the consequences of offloading on subsequent performance within a complex task, these effects were assessed only in the short term. Currently, little is known about the long-term consequences of offloading, largely due to the challenge of examining long-term effects within a controlled environment beyond the laboratory setting. The literature presents mixed findings, with a notable limitation being that most are correlational studies. Some studies highlight the potential negative effects of offloading. For example, research on GPS usage among regular drivers has suggested its adverse impacts on spatial memory, as GPS reliance may decrease engagement with the environment, reducing hippocampal activity (Dahmani & Bohbot, 2020). Conversely, other studies indicate that using digital aids over time may contribute to a "digital reserve", supporting cognitive function, especially in older age (Scullin et al., 2024; Benge & Scullin, 2023). The available evidence, however, is still scarce and does not allow for any clear conclusion. Therefore, more research is needed on the long-term consequences of different kinds of offloading in different cognitive domains.

Taken together, our studies contributed to the growing body of knowledge on cognitive offloading by describing how and when cognitive offloading occurs in the previously uninvestigated domain of route planning on maps, and observing how the opportunity to use offloading changes the cognitive strategies employed to carry out the planning task and the perception of the task. We also highlighted the advantages and the drawbacks of the use of offloading in this domain and the transfer effects related to the availability/unavailability of offloading. Finally, we investigated how visual aids for offloading changed performance and task perception, suggesting potential applied developments. Overall, our results offer a novel theoretical and empirical contribution to the research on cognitive offloading, hopefully suggesting promising research directions.

References

Allain, P., Nicoleau, S., Pinon, K., Etcharry-Bouyx, F., Barré, J., Berrut, G., ... & Le Gall, D. (2005). Executive functioning in normal aging: A study of action planning using the Zoo Map Test. *Brain and Cognition*, *57*(1), 4-7. https://doi.org/10.1016/j.bandc.2004.08.011

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science*, *26*(1), 39-83. https://doi.org/10.1207/s15516709cog2601_2

Armitage, K. L., Bulley, A., & Redshaw, J. (2020). Developmental origins of cognitive offloading. *Proceedings of the Royal Society B*, *287*, 20192927. http://doi.org/10.1098/rspb.2019.2927.

Arthurs, L. A., Baumann, S. P., Rice, J. M., & Litton, S. D. (2021). The development of individuals' map-reading skill: what research and theory tell us. *International Journal of Cartography*, *9*(1), 3–28. https://doi.org/10.1080/23729333.2021.1950318

Bailenson, J. N., Shum, M. S., & Uttal, D. H. (1998). Road climbing: Principles governing asymmetric route choices on maps. *Journal of Environmental Psychology*, *18*(3), 251-264. https://doi.org/10.1006/jevp.1998.0095

Basso, D., Bisiacchi. P. S., Cotelli, M., & Farinello, C. (2001). Planning times during travelling salesman's problem: Differences between closed head injury and normal subjects. *Brain and Cognition, 46,* 38–42. https://doi.org/10.1016/S0278-2626(01)80029-4

Benge, J., & Scullin, M. K. (2023). The technological reserve hypothesis: A meta-analysis. *Alzheimer's & Dementia*, *19*, e071658. https://doi.org/10.1002/alz.071658

Bilalić, M., & McLeod, P. (2014). Why Good Thoughts Block Better Ones. *Scientific American*, *310*(3), 74–79. https://www.jstor.org/stable/26039821

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, *108*(3), 652-661. https://doi.org/10.1016/j.cognition.2008.05.005

Bilalić, M., McLeod, P., & Gobet, F. (2008a). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition, 108*(3), 652–661. https://doi.org/10.1016/j.cognition.2008.05.005

Bilalić, M., McLeod, P., & Gobet, F. (2008b). Inflexibility of experts—reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology, 56*(2), 73–102. https://doi.org/10.1016/j.cogpsych.2007.02.001

Bilalić, M., McLeod, P., & Gobet, F. (2010). The mechanism of the Einstellung (set) effect: A pervasive source of cognitive bias. *Current Directions in Psychological Science, 19*(2), 111–115. https://doi.org/10.1177/0963721410363571

Bluestein, N., & Acredolo, L. (1979). Developmental changes in map-reading skills. Child Development, 691–697. https://doi.org/10.2307/1128934

Boot, W. R., Becic, E., & Kramer, A. F. (2009). Stable individual differences in search strategy?: The effect of task demands and motivational factors on scanning strategy in visual search. *Journal of Vision*, *9*(3):7, 1–16. https://doi.org/10.1167/9.3.7

Brügger, A., Richter, K. F., & Fabrikant, S. I. (2019). How does navigation system behavior influence human behavior?. *Cognitive Research: Principles and Implications*, *4*(5), 1-22. https://doi.org/10.1186/s41235-019-0156-5

Burns, N. R., Lee, M. D., & Vickers, D. (2006). Are individual differences in performance on perceptual and cognitive optimization problems determined by general intelligence?. The *Journal of Problem Solving*, *1*(1), 5-19. http://dx.doi.org/10.7771/1932-6246.1003

Carlson, R. A., Avraamides, M. N., Cary, M., & Strasberg, S. (2007). What do the hands externalize in simple arithmetic? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 747–756. https://doi.org/10.1037/0278-7393.33.4.747

Carr, N. (2008). Is Google making us stupid?. *Teachers College Record*, *110*(14), 89-94. https://doi.org/10.1177/016146810811001427

Chronicle, E. P., MacGregor, J. N., Lee, M., Ormerod, T. C., & Hughes, P. (2008). Individual differences in optimization problem solving: Reconciling conflicting results. *The Journal of Problem Solving*, *2*(1), 41-49. http://dx.doi.org/10.7771/1932-6246.1030

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Coupey, E. (1994). Restructuring: constructive processing of information displays in consumer choice. *Journal of Consumer Research*, *21*(1), 83-99. https://doi.org/10.1086/209384

Dahmani, L., & Bohbot, V. D. (2020). Habitual use of GPS negatively impacts spatial memory during self-guided navigation. *Scientific Reports*, *10*(1), 6310. https://doi.org/10.1038/s41598-020-62877-0

Dalton, R. C., Hölscher, C., & Montello, D. R. (2019). Wayfinding as a social activity. *Frontiers in psychology*, *10*, 142. https://doi.org/10.3389/fpsyg.2019.00142

Del Missier, F., Sassano, A., Coni, V., Salomonsson, M., & Mäntylä, T. (2018) Blocked vs. interleaved presentation and proactive interference in episodic memory, *Memory*, *26*(5), 697-711. https://doi.org/10.1080/09658211.2017.1402937

Del Missier, F., Stragà, M., Visentini, M., Munaretto, G., & Mäntylä, T. (2021). Intention Framing in Time-Based Prospective Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(8), 1226-1245. https://doi.org/10.1037/xlm0000991

Dobson, M. W. (1977). Eye movement parameters and map reading. *The American Cartographer, 4*(1), 39–58. https://doi.org/10.1559/152304077784080022

Dunn, T. L., & Risko, E. F. (2016). Toward a metacognitive account of cognitive offloading. *Cognitive Science*, *40*(5), 1080-1127.

Engeler, N. C., & Gilbert, S. J. (2020). The effect of metacognitive training on confidence and strategic reminder setting. *PLoS One*, *15*(10), e0240858. https://doi.org/10.1371/journal.pone.0240858

Epstein, R. A., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017). The cognitive map in humans: spatial navigation and beyond. *Nature Neuroscience*, *20*(11), 1504-1513. https://doi.org/10.1038/nn.4656

Eskritt, M., & Ma, S. (2014) Intentional forgetting: note-taking as a naturalistic example. *Memory & Cognition*, *42*, 237–246. https://doi.org/10.3758/s13421-013-0362-1

Fenech, E. P., Drews, F. A., & Bakdash, J. Z. (2010, September). The effects of acoustic turn-by-turn navigation on wayfinding. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 54, No. 23, pp. 1926-1930). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/154193121005402

Fenech, E. P., Drews, F. A., & Bakdash, J. Z. (2010, September). The effects of acoustic turn-by-turn navigation on wayfinding. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 54, No. 23, pp. 1926-1930). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/154193121005402305

Finkler, A. L., & Meyerhoff, H. S. (2024, September 11-13). *Intentional forgetting prevents proactive interference, but does not improve memory for novel stimuli* [Conference session]. Cognitive Offloading Meeting 2024, Geneve, Switzerland. https://mattermost.samgilbert.net/conference2024/

Finkler, A. L., Meyerhoff, H. S. (2024, September 11-13). *Intentional forgetting prevents proactive interference, but does not improve memory for novel stimuli* [Conference session]. Cognitive Offloading Meeting 2024, Geneve, Switzerland. https://mattermost.samgilbert.net/conference2024/

Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the Internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General*, *144*(3), 674–687. https://doi.org/10.1037/xge0000070

Fröscher, L., Friedrich, A.-K., Berentelg, M., Widmer, C., Gilbert, S. J., & Papenmeier, F. (2022). Framing cognitive offloading in terms of gains or losses: Achieving a more optimal use of reminders. *Cognitive Research: Principles and Implications*, *7*(61). https://doi.org/10.1186/s41235-022-00416-3

Fum, D., & del Missier, F. (2000a). Climbing the mazes: A cognitive model of spatial planning. In *Proceedings of the third international conference on cognitive modeling* (pp. 126-133). Universal Press, Veenendaal.

Fum, D., & Del Missier, F. (2000b, August 5-7). *Adaptive spatial planning: An ACT-R model*. Seventh ACT-R Annual Workshop, Carnegie Mellon University, Pittsburgh, USA.

Gardony, A. L., Brunyé, T. T., & Taylor, H. A. (2015) Navigational aids and spatial memory impairment: the role of divided attention. *Spatial Cognition & Computation*, *15*(4), 246–284. https://doi.org/10.1080/13875868.2015.1059432

Gardony, A. L., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2013). How navigational aids impair spatial memory: Evidence for divided attention. *Spatial Cognition & Computation*, *13*(4), 319-350. https://doi.org/10.1080/13875868.2013.792821

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. J*ournal of Experimental Psychology: General*, *112*(1), 58–72. https://doi.org/10.1037/0096-3445.112.1.58

Gilbert, S. J. (2015a). Strategic offloading of delayed intentions into the external environment. *The Quarterly Journal of Experimental Psychology*, *68*(5), 971–992. https://doi.org/10.1126/science.1207745

Gilbert, S. J. (2015b). 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. https://doi.org/10.1016/j.concog.2015.01.006

Gilbert, S. J., Bird, A., Carpenter, J. M., Fleming, S. M., Sachdeva, C., & Tsai, P. C. (2020). Optimal use of reminders: Metacognition, effort, and cognitive offloading. *Journal of Experimental Psychology: General*, *149*(3), 501–517. https://doi.org/10.1037/xge0000652

Gilbert, S. J., Boldt, A., Sachdeva, C., Scarampi, C., & Tsai, P. C. (2023). Outsourcing memory to external tools: A review of 'intention offloading'. *Psychonomic Bulletin & Review*, *30*(1), 60-76. https://doi.org/10.3758/s13423-022-02139-4

Gilbert, S., Benett, A., & O'Brien, K. (2024, September 11-13). *Metacognition guides optimal cognitive offloading with ChatGPT* [Conference session]. Cognitive Offloading Meeting 2024, Geneve, Switzerland. https://mattermost.samgilbert.net/conference2024/

Grinschgl, S., & Venus, C. (2024, September 11-13). *Cognitive offloading in the lab and in daily life* [Conference session]. Cognitive Offloading Meeting 2024, Geneve, Switzerland. https://mattermost.samgilbert.net/conference2024/

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. https://doi.org/10.1016/j.chb.2020.106317

Grinschgl, S., Papenmeier, F., & Meyerhoff, H. S. (2021). Consequences of cognitive offloading: Boosting performance but diminishing memory. *Quarterly Journal of Experimental Psychology*, *74*(9), 1477-1496. https://doi.org/10.1177/17470218211008060

Guthrie, L. G., Vallée-Tourangeau, F., Vallée-Tourangeau, G., & Howard, C. (2015). Learning and interactivity in solving a transformation problem. *Memory & cognition*, *43*, 723-735.

Hambrick, D. Z., & Altmann, E. M. (2015). The role of placekeeping ability in fluid intelligence. *Psychonomic Bulletin & Review*, *22*, 1104–1110. https://doi.org/10.3758/s13423-014-0764-5

Harris, J. E. (1980). Memory aids people use: Two interview studies. *Memory & Cognition*, *8*(1), 31-38. https://doi.org/10.3758/BF03197549

Hayes-Roth, B. (1979). *Flexibility in executive strategies*. N:1170, The Rand Corporation, Santa Monica, California.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science*, *3*(4), 275-310. https://doi.org/10.1016/S0364-0213(79)80010-5

Hayes-Roth, B., & Thorndyke, P. (1979). *Decision-making during the planning process*. N:1213, The Rand Corporation, Santa Monica, California.

Henkel, L. A. (2014). Point-and-shoot memories: The influence of taking photos on memory for a museum tour. *Psychological Science*, *25*(2), 396-402. https://doi.org/10.1177/0956797613504438

Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2014). Pillars of judgment: How memory abilities affect performance in rule-based and exemplar-based judgments. *Journal of Experimental Psychology: General*, *143*(6), 2242–2261. https://doi.org/10.1037/a0037989

Hölscher, C., Tenbrink, T., & Wiener, J. M. (2011). Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition*, *121*(2), 228-247.

Holt, D. V., Wolf, J., Funke, J., Weisbrod, M., & Kaiser, S. (2013). Planning impairments in schizophrenia: Specificity, task independence and functional relevance. *Schizophrenia Research*, *149*(1-3), 174-179. https://doi.org/10.1016/j.schres.2013.06.018

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science*, *19*(3), 265-288. https://doi.org/10.1207/s15516709cog1903_1

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of experimental psychology: General*, *133*(2), 189.

Kirsh, D., & and Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. https://doi.org/10.1016/0364-0213(94)90007-8

Kong, X., & Schunn, C. D. (2007). Global vs. local information processing in visual/spatial problem solving: The case of traveling salesman problem. *Cognitive Systems Research*, *8*(3), 192-207. https://doi.org/10.1016/j.cogsys.2007.06.002

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.

Liu, J., Singh, A. K., Wunderlich, A., Gramann, K., & Lin, C. T. (2022). Redesigning navigational aids using virtual global landmarks to improve spatial knowledge retrieval. npj Science of Learning, 7, 17. https://doi.org/10.1038/s41539-022-00132-z

Loess, H. (1968). Short-term memory and item similarity. *Journal of Verbal Learning and Verbal Behavior*, *7*(1), 87-92. https://doi.org/10.1016/S0022-5371(68)80169-0

Luchins, A. S., & Luchins, E. H. (1950). New experimental attempts at preventing mechanization in problem solving. *The Journal of General Psychology*, *42*(2), 279-297. https://doi.org/10.1080/00221309.1950.9920160

MacGregor, J. N., & Chu, Y. (2011). Human performance on the traveling salesman and related problems: A review. *The Journal of Problem Solving*, *3*(2), 1-29.

MacGregor, J. N., & Ormerod, T. (1996). Human performance on the traveling salesman problem. *Perception & Psychophysics*, *58*, 527-539. https://doi.org/10.3758/BF03213088

Mandel, N., & Johnson, E. J. (2002). When web pages influence choice: Effects of visual primes on experts and novices. *Journal of Consumer Research*, *29*(2), 235–245. https://doi.org/10.1086/341573

Marra, C., Gainotti, G., Scaricamazza, E., Piccininni, C., Ferraccioli, M., & Quaranta, D. (2013). The Multiple Features Target Cancellation (MFTC): an attentional visual conjunction search test. Normative values for the Italian population. *Neurological Sciences*, *34*, 173-180. https://doi.org/10.1007/s10072-012-0975-3

McDonald, Z. C., & Soares, J. S. (2024, September 11-13). *Digital fixation in the age of AI: Examining mental set and cognitive style as digital fixation mechanisms* [Conference session]. Cognitive Offloading Meeting 2024, Geneve, Switzerland. https://mattermost.samgilbert.net/conference2024/

Meyerhoff, H. S., Grinschgl, S., Papenmeier, F., & Gilbert, S. J. (2021). Individual differences in cognitive offloading: A comparison of intention offloading, pattern copy, and short-term memory capacity. *Cognitive Research: Principles and Implications*, *6*(1), 34. https://doi.org/10.1186/s41235-021-00298-x

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. *Neurological Sciences*, *34*, 749-754. http:// doi.org/10.1007/s10072-012-1130-x

Montello, D. R., and Sas, C. (2006). Human factors of wayfinding in navigation," in *International Encyclopedia of Ergonomics and Human Factors*, 2nd Edn, ed. W. Karwowski (London: CRC Press/Taylor & Francis, Ltd), 2003–2008. doi: 10.1111/j.1756-8765.2009.01028.x

Moritz, J., Meyerhoff, H. S., Meyer-Dernbecher, C., & Schwan, S. (2018). Representation control increases task efficiency in complex graphical representations. *PLoS ONE*, *13*(4), e0196420. https://doi.org/10.1371/journal.pone.0196420

Morris, R., & Ward, G. (Eds.). (2004). *The cognitive psychology of planning*. Psychology Press.

Muffato, V., & Meneghetti, C. (2020). Knowledge of familiar environments: Assessing modalities and individual visuo-spatial factors. *Journal of Environmental Psychology*, *67*, 101387.

Muffato, V., Feraco, T., Miola, L., Tortora, C., Pazzaglia, F., & Meneghetti, C. (2021). Finding the shortest path in a familiar environment: A comparison between describing and walking a path after accounting for the role of individual factors. *Journal of Environmental Psychology*, *78*, 101708.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-Hall.

Norman, D. (1991). Cognitive artifacts. In J. M. Carroll (Ed.), *Designing interaction: Psychology at the human-computer interface* (pp. 17–38). Cambridge University Press.

O'Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive Psychology*, *35*(1), 34-70. https://doi.org/10.1006/cogp.1997.0676

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. Cambridge University Press.

Oosterman, J. M., Wijers, M., & Kessels, R. P. (2013). Planning or something else? Examining neuropsychological predictors of Zoo Map performance. *Applied Neuropsychology: Adult*, *20*(2), 103-109. https://doi.org/10.1080/09084282.2012.670150

Ormerod, T. C., & Chronicle, E. P. (1999). Global perceptual processing in problem solving: The case of the traveling salesperson. *Perception & Psychophysics*, *61*, 1227-1238. https://doi.org/10.3758/BF03207625

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, *47*, 1343-1355. https://doi.org/10.3758/s13428-014-0543-2

Phillips, L. H., Wynn, V. E., McPherson, S., & Gilhooly, K. J. (2001). Mental planning and the Tower of London task. *The Quarterly Journal of Experimental Psychology Section A*, *54*(2), 579-597.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*.

Reitter, D., & Lebiere, C. (2010). A cognitive model of spatial path-planning. *Computational and Mathematical Organization Theory*, *16*, 220-245. https://doi.org/10.1007/s10588-010-9073-3

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676-688. https://doi.org/10.1016/j.tics.2016.07.002

Risko, E. F., Medimorec, S., Chisholm, J. and Kingstone, A. (2014). Rotating with rotated text: A natural behavior approach to investigating cognitive offloading. *Cognitive Science*, *38*, 537-564. https://doi.org/10.1111/cogs.12087

Sachdeva, C., & Gilbert, S. J. (2020). Excessive use of reminders: Metacognition and effort-minimisation in cognitive offloading. *Consciousness and Cognition*, *85*, 103024. https://doi.org/10.1016/j.concog.2020.103024

Scarampi, C., & Gilbert, S. J. (2020). The effect of recent reminder setting on subsequent strategy and performance in a prospective memory task. *Memory*, *28*(5), 677–691. https://doi.org/10.1080/09658211.2020.1764974

Scarampi, C., & Gilbert, S. J. (2021). Age differences in strategic reminder setting and the compensatory role of metacognition. *Psychology and Aging*, *36*(2), 172. https://doi.org/10.1037/pag0000590

Schönpflug, W. (1986). The trade-off between internal and external information storage. *Journal of Memory and Language*, *25*(6), 657-675. https://doi.org/10.1016/0749-596X(86)90042-2

Scullin, M. K., Benge, J. F., Tindaan, R., & David, K. (2024, September 11-13). *The technological reserve hypothesis: Smartphone- versus paper-based intention offloading in mild cognitive impairment and dementia* [Conference session]. Cognitive Offloading Meeting 2024, Geneve, Switzerland. https://mattermost.samgilbert.net/conference2024/

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4–27. https://doi.org/10.1037/0096-3445.125.1.4

Shelton, A. L., & Gabrieli, J. D. E. (2002). Neural correlates of encoding space from route and survey perspectives. *Journal of Neuroscience*, *22*(7), 2711–2717. https://doi.org/10.1523/JNEUROSCI.22-07-02711.2002

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, *7*(2), 268-288. https://doi.org/10.1016/0010-0285(75)90012-2

Soares, J. S., & Storm, B. C. (2018). Forget in a flash: A further investigation of the photo-taking-impairment effect. *Journal of Applied Research in Memory and Cognition*, *7*(1), 154-160. https://doi.org/10.1016/j.jarmac.2017.10.004

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, *333*(6043), 776-778. https://doi/10.1126/science.1207745

Storm, B. C., & Stone, S. M. (2015). Saving-enhanced memory: The benefits of saving on the learning and remembering of new information. *Psychological Science*, *26*(2), 182-188. https://doi.org/10.1177/0956797614559285

Thomas, C., Didierjean, A., & Kuhn, G. (2018). It is magic! How impossible solutions prevent the discovery of obvious ones? *Quarterly Journal of Experimental Psychology, 71*(12), 2481–2487. https://doi.org/10.1177/1747021817743439

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, *14*(4), 560-589. https://doi.org/10.1016/0010-0285(82)90019-6

van Rooij, I., Schactman, A., Kadlec, H., & Stege, U. (2006). Perceptual or analytical processing?: evidence from children's and adult's performance on the Euclidean Traveling Salesperson problem. *Journal of Problem Solving*, *1*(1), 44-73. https://doi.org/10.7771/1932-6246.1006

Vickers, D., Butavicius, M., Lee, M. D., & Medvedev, A. (2001). Human performance on visually presented Traveling Salesman Problems. *Psychological Research*, *65*, 34–45. https://doi.org/10.1007/s004260000031

Vickers, D., Mayo, T., Heitmann, M., Lee, M. D., & Hughes, P. (2004). Intelligence and individual differences on three types of visually presented optimisation problems. *Personality and Individual Differences*, *36*(5), 1059–1071. https://doi.org/10.1016/S0191-8869(03)00200-9

Ward, A. F. (2021). People mistake the internet's knowledge for their own. *Proceedings of the National Academy of Sciences*, *118*(43), e2105061118. https://doi.org/10.1073/pnas.2105061118

Ward, A. F. (2021). People mistake the internet's knowledge for their own. *Proceedings of the National Academy of Sciences*, *118*(43), e2105061118. https://doi.org/10.1073/pnas.2105061118

Weis, P. P., & Wiese, E. (2018). Using tools to help us think: Actual but also believed reliability modulates cognitive offloading. *Human Factors*, *61*, 243–254. https://doi.org/10.1177/0018720818797553

Weis, P. P., & Wiese, E. (2019). Problem solvers adjust cognitive offloading based on performance goals. *Cognitive Science*, *43*, e12802. https://doi.org/10.1111/cogs.12802

Wiener, J. M., & Mallot, H. A. (2003). 'Fine-to-coarse'route planning and navigation in regionalized environments. *Spatial cognition and computation*, *3*(4), 331-358.

Wiener, J. M., Büchner, S. J., & Hölscher, C. (2009). Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition & Computation*, *9*(2), 152-165. https://doi.org/10.1080/13875860902906496

Wiener, J. M., Ehbauer, N. N., & Mallot, H. A. (2009). Planning paths to multiple targets: memory involvement and planning heuristics in spatial problem solving. *Psychological Research PRPF*, *73*(5), 644-658.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome*. St Edmunds, UK: Thames Valley Test Company.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 0058. https://doi.org/10.1038/s41562-017-0058

Woo, S. E., Chernyshenko, O. S., Longley, A., Zhang, Z. X., Chiu, C. Y., & Stark, S. E. (2014). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of personality assessment*, *96*(1), 29-45. https://doi.org/10.1080/00223891.2013.806328

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179-217. https://doi.org/10.1016/S0364-0213(99)80022-6

APPENDIX A

1. Coding Scheme for Offloading Strategies

The coding scheme was developed before data collection. For this purpose, we relied on an *a-priori* cognitive task analysis and on preliminary testing, which led us to expect observing four main types of external offloading strategies (see Figure 2 in the main text): (1) marking out start and arrival locations; (2) marking out intermediate locations to be visited; (3) marking out ordering constraints; (4) track keeping. In the main study, various specific implementations of these four types of offloading strategies were observed, but no alternative offloading strategies were identified.

The coding scheme and the observed implementations of the four expected strategies are presented in Table A.A.1. The use of each of the four types of offloading strategies was assessed for each of the maps completed by participants in the offloading condition, as specified in the description of the main study. Specifically, each type of offloading strategy was assigned a score of one when it was applied in the map through at least one of its implementations, and a score of zero when it was not employed. Moreover, each map completed by each participant in the offloading condition received a summative score ranging from zero (no offloading strategy employed) to four (all offloading strategies employed).

Table A.A.1

Coding Scheme for Offloading Strategies

Offloading Strategies	Observed implementations	Description of the observed implementations
Marking out starting and arrival locations	Starting and arrival locations marked out in the same way	Starting and arrival locations marked out differently from the other locations but not from each other.
	Starting and arrival locations marked out differently	Starting and arrival locations marked out differently both from the other locations in the route and from each other.
Marking out intermediate locations to be visited	Intermediate locations to be visited marked out	Dashes, dots, or other symbols drawn near the locations to be visited, making them visually distinguishable from the distractor locations.
Marking out ordering constraints	Locations constrained by order of visit marked out with symbols	Symbols drawn near the order-constrained locations (e.g., 1 near the first place to be visited and 2 near the second one; stars drawn near two constrained locations, circles drawn near two other constrained locations).
	Locations constrained by order of visit connected with arrow(s)	Arrow connecting the first location to be visited with the second one.
	Locations constrained by order of visit marked out in the instructions frame of the map	Order-constrained locations highlighted in the rightmost panel providing specific instructions for the map.
	Locations constrained by order of visit written as in the solution format outside the response frame	Pair of letters corresponding to order-constrained locations written on the map sheet in the same format as it needed to be reported in the response frame.
Track keeping	Planned route tracked	Path connecting all the locations to be visited drawn on the map.
	Planned route numbered	Locations to be visited numbered on the map and/or in the instructions frame, in accordance with their sequential position in the planned route
	Planned route written as in the solution format outside the response frame	Planned route written on the map sheet in the same format as it needed to be reported in the response frame.
	Locations to be visited marked out in the instruction frame	Locations to be included in the route marked out in the instructions frame with the same symbol, to keep track of the locations to be visited (e.g., for checking the solution)

Table A.A.2

Descriptive Statistics for Gender in Offloading and No Offloading Conditions

	Offloading	No offloading	χ^2	df	p
F	43	48	1.75	1	0.186
M	11	6			

Note. One participant in the offloading condition was excluded from the analysis due to their preference not to indicate their gender.

Table A.A.3

Descriptive Statistics for Visual Search and Spatial Working Memory in the Offloading and No Offloading Conditions

	Offloading	No offloading	Student's t	df	p	95%CI Lower	95%CI Upper	Cohen's d
Visual Search (MFTC)	0.90 (0.08)	0.89 (0.010)	-0.8	107	0.425	-0.05	0.02	-0.15
Working Memory (Symmetry Span)	27.78 (7.99)	26.54 (7.14)	-0.86	107	0.393	-4.12	1.63	-0.16

Note. Lower and upper bounds for 95% confidence intervals are indicated with 95%CI Lower and 95%CI Upper, respectively.

Table A.A.4

Correlations Between All Performance Measures in the Offloading Condition

	1	2	3	4	5	6	7	8	9	10
1. Proportion of high difficulty maps without errors	—									
2. Proportion of low difficulty maps without errors	0.10	—								
3. Completion time for high difficulty maps	0	0.08	—							
4. Completion time for low difficulty maps	0.07	0.12	0.65***	—						
5. Solution time for high difficulty maps	-0.09	0.23	0.26	0.15	—					
6. Solution time for low difficulty maps	-0.18	0.08	0.24	0.29*	0.38**	—				
7. Proportional deviation for high difficulty maps	0.01	0.11	-0.2	-0.15	0.51***	0.36**	—			
8. Proportional deviation for low difficulty maps	-0.26	0.22	-0.11	0.09	0.18	0.4**	0.19	—		
9. Visual Search (MFTC)	-0.03	0.13	0.18	0.02	-0.25	-0.11	-0.44***	0.12	—	
10. Working Memory (Symmetry Span)	0.33*	-0.06	-0.19	-0.23	-0.17	-0.31*	-0.23	-0.05	0.16	—

*$p < .05$. **$p < .01$. ***$p < .001$

Table A.A.5

Correlations Between all Performance Measures in the No Offloading Condition

	1	2	3	4	5	6	7	8	9	10
1. Proportion of high difficulty maps without errors	—									
2. Proportion of low difficulty maps without errors	0.13	—								
3. Completion time for high difficulty maps	0.23	-0.25	—							
4. Completion time for low difficulty maps	0	-0.07	-0.05	—						
5. Solution time for high difficulty maps	-0.13	-0.17	0.07	0.01	—					
6. Solution time for low difficulty maps	-0.09	0.06	-0.08	0.34*	0.37	—				
7. Proportional deviation for high difficulty maps	-0.69***	0.11	-0.38**	0.22	0.07	0.23	—			
8. Proportional deviation for low difficulty maps	-0.12	-0.07	-0.17	0.41**	0.02	-0.03	0.3*	—		
9. Visual Search (MFTC)	0.24	0.01	0.11	0.14	-0.03	-0.06	-0.19	-0.09	—	
10. Working Memory (Symmetry Span)	0.33*	0.23	-0.11	-0.09	-0.21	-0.08	-0.33*	0.00	0.12	—

*$p < .05$. **$p < .01$. ***$p < .001$

Table A.A.6

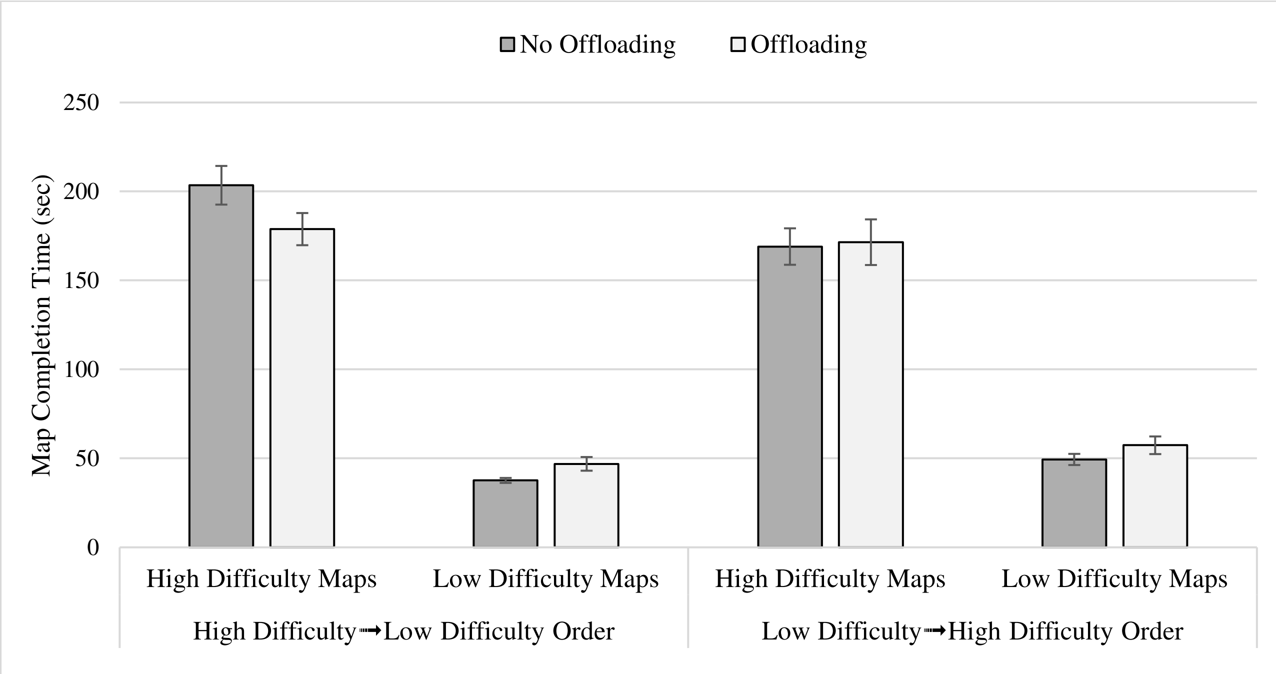
Further Analyses on Map Completion Time: Descriptive Statistics and Independent Sample T-Tests Within Each Map Difficulty Presentation Order

	High→Low Difficulty Order					**Low→High Difficulty Order**				
	Offloading M (SD)	No Offloading M (SD)	t	p	Cohen's d	Offloading M (SD)	No Offloading M (SD)	t	p	Cohen's d
High Difficulty Maps	178.83s (46.07s)	203.47s (55.40s)	1.74	.044	.48	171.49s (67.89s)	169.03s (52.23s)	-0.15	.559	-.04
Low Difficulty Maps	46.90s (20.01s)	37.60s (7.36s)	-2.27	.030	.62	57.36s (26.23s)	49.38s (15.97s)	-1.36	.181	-.37

Note. **High difficulty maps**: one-tailed t-tests were conducted, as we expected participants in the Offloading condition to spend less time than those in the No Offloading condition to complete these maps (H2); t refers to Student's t. **Low difficulty maps**: two-tailed t-tests were conducted, as we did not expect to find differences between participants in the Offloading vs. No Offloading condition; t refers to Welch's t, as Levene's test indicated a violation of the equal variance assumption (when considering Student's t, results do not significantly change).

Figure A.A.1

Further Analyses on Map Completion Time: Map completion Time as a Function of Map Difficulty and Presentation Order



Note. Bars represent standard errors.

Table A.A.7

Multiple Regression on the Proportion of Maps with Starting and Arrival Places Marked Out in the Offloading Condition

	High difficulty maps					**Low difficulty maps**				
Predictors	β	95%CI Lower	95%CI Upper	t	p	β	95%CI Lower	95%CI Upper	t	p
Visual Search (MFTC)	0.37	0.09	0.66	2.63	0.011	0.17	-0.09	0.42	1.33	0.190
Working Memory (Symmetry Span)	0	-0.26	0.27	0.03	0.975	0.01	-0.25	0.26	0.04	0.965
Metacognitive prediction	-0.08	-0.37	0.21	-0.57	0.571	-0.03	-0.28	0.23	-0.21	0.838
Presentation order of map difficulty	-0.45	-0.98	0.08	-1.7	0.095	-0.9	-1.41	-0.4	-3.62	< .001
Overall Model	Adjusted R²		F		p	Adjusted R²		F		p
	0.10		2.53		0.052	0.17		3.84		0.009

Table A.A.8

Multiple Regression on the Proportion of Maps with Intermediate Locations to be Visited Marked Out in the Offloading Condition

	High difficulty maps					Low difficulty maps				
Predictors	β	95%CI Lower	95%CI Upper	t	p	β	95%CI Lower	95%CI Upper	t	p
Visual Search (MFTC)	0.04	-0.26	0.34	0.24	0.808	0.17	-0.1	0.44	1.24	0.220
Working Memory (Symmetry Span)	0.23	-0.05	0.51	1.68	0.100	0.11	-0.16	0.39	0.81	0.419
Metacognitive prediction	0.06	-0.25	0.36	0.37	0.716	-0.12	-0.39	0.16	-0.85	0.399
Presentation order of map difficulty	0.01	-0.55	0.57	0.03	0.972	-0.5	-1.05	0.04	-1.87	0.067
Overall Model	Adjusted R²		F		p	Adjusted R²		F		p
	-0.01		0.85		0.499	0.04		1.5		0.216

Table A.A.9

Multiple Regression on the Proportion of Maps with Ordering Constraints Marked Out in the Offloading Condition

	High difficulty maps					Low difficulty maps				
Predictors	β	95%CI Lower	95%CI Upper	t	p	β	95%CI Lower	95%CI Upper	t	p
Visual Search (MFTC)	0.19	-0.09	0.47	1.36	0.179	0.06	-0.17	0.29	0.51	0.609
Working Memory (Symmetry Span)	0.23	-0.04	0.49	1.74	0.089	0.28	0.05	0.51	2.42	0.019
Metacognitive prediction	0.08	-0.2	0.37	0.59	0.557	0.08	-0.15	0.31	0.7	0.489
Presentation order of map difficulty	-0.42	-0.94	0.11	-1.59	0.117	-1.03	-1.48	-0.58	-4.56	< .001
Overall Model	Adjusted R²		F		p	Adjusted R²		F		p
	0.11		2.62		0.046	0.32		7.49		< .001

Table A.A.10

Multiple Regression on the Proportion of Maps with Externally Tracked Task Progression in the Offloading Condition

	High difficulty maps					Low difficulty maps				
Predictors	β	95%CI Lower	95%CI Upper	t	p	β	95%CI Lower	95%CI Upper	t	p
Visual Search (MFTC)	0.31	0.01	0.6	2.08	0.043	0.15	-0.12	0.42	1.1	0.277
Working Memory (Symmetry Span)	-0.08	-0.36	0.19	-0.62	0.538	-0.18	-0.46	0.09	-1.35	0.183
Metacognitive prediction	-0.2	-0.5	0.1	-1.36	0.18	0	-0.28	0.27	-0.02	0.988
Presentation order of map difficulty	-0.18	-0.73	0.38	-0.64	0.526	-0.53	-1.06	0.01	-1.97	0.054
Overall Model	Adjusted R²		F		p	Adjusted R²		F		p
	0.02		1.24		0.306	0.05		1.74		0.155

Table A.A.11

Multiple Regression On the Proportional Deviation Above the Shortest Path in the No Offloading Condition

	High difficulty maps					Low difficulty maps				
Predictors	β	95%CI Lower	95%CI Upper	t	p	β	95%CI Lower	95%CI Upper	t	p
Visual Search (MFTC)	-.13	-.40	.14	-0.99	.320	-.10	-.39	.19	-0.71	.482
Working Memory (Symmetry Span)	-.33	-.59	-.06	-2.46	.018	.01	-.27	.03	.10	.918
Predictive metacognitive judgments	-.17	-.44	.10	-1.24	.220	.05	-.24	.33	0.32	.749
Overall Model	Adjusted R²		F		p	Adjusted R²		F		p
	.11		3.13		.034	-.05		0.18		.907

In the high difficulty maps, only the Symmetry Span score was a significant predictor of the proportional deviation above the shortest route. In the low difficulty maps, none of the variables included in the regression model was a significant predictor of the proportional deviation above the shortest path.

APPENDIX B

1. Study 2 – Analysis for the Training Block

In the training block of Study 2, participants completed the route planning task under one of three between-participants conditions (pen, no-pen, and no-hand), without prior experience of the task. Therefore, in the following analyses, we included the training condition as a between-participants factor, rather than the group factor used in the test blocks.

Table A.B.1

One-Way ANOVA for the Effect of Training Condition on the Proportion of Maps Completed Without Errors in the Training Block

	SS	df	MS	F	p	η^2p
Training condition	0.04	2	0.02	1.06	.350	0.02
Residuals	2.28	134	0.02			

Table A.B.2

One-Way ANOVA for the effect of Training Condition on the Proportion of maps completed without errors in the Training Block

	SS	df	MS	F	p	η^2p
Training	0	2	0.00	0.42	.655	0.01
Residuals	0.61	134	0.00			

Table A.B.3

One-Way ANOVA for the Effect of Training Condition on Map Completion Time

	SS	df	MS	F	p	η^2p
Training	34814.74	2	17407.37	4.35	0.015	0.06
Residuals	536444.19	134	4003.31			
Post-hoc comparisons	Mean Difference	SE	df	t	p	Cohen's d
pen vs. no-pen	-39.25	13.33	134	-2.94	0.011	-0.62
pen vs. no-hand	-10.46	13.33	134	-0.78	0.713	-0.17
no-pen vs. no-hand	28.79	15.58	134	1.85	0.158	0.46

Note. Post-hoc comparisons were conducted using Tukey's test.

2. Study 2 - Correlations Between Individual Differences in Cognitive Abilities and Route Planning Task With/Without Offloading

A series of analyses were conducted to assess whether the correlations with individual differences observed in the Study 2 aligned with the patterns identified in Study 1. Although we had four different groups (pen→no-pen, pen→no-hand, no-pen→pen, no-hand→pen), the route planning task could only be performed under three conditions: one allowing for the use of offloading strategies (pen condition) and two not allowing them (no-pen and no-hand). To enhance statistical power, the correlational analyses were conducted for each condition rather than for each group.

We first report the correlations between individual differences in cognitive abilities and route planning performance in the test blocks. Given that the training condition could influence performance in the test blocks, we also report the correlations between individual differences and route planning performance in the training block, where participants had no prior experience with the task.

Participants with better visual search abilities were expected to use more offloading strategies when planning routes in the pen condition. Furthermore, greater use of offloading strategies was expected to be associated with greater route planning accuracy, reflected in shorter deviations from the shortest route.

The correlation between visual search abilities (MFTC score) and the number of offloading strategies used in the pen test block did not reach statistical significance, $r(135) = .15$, $p = .089$. However, in line with H4a, the greater the number of offloading strategies used, the smaller the deviation above the shortest route, $r(134) = -.23$, $p = .007$. Notably, a lower deviation above the shortest route was also correlated with higher spatial working memory, $r(134) = -.23$, $p = .006$.

In the training block, better visual search abilities were associated with a greater use of offloading strategies, $r(69) = .32$, $p = .006$. The number of offloading strategies was negatively correlated with the deviation above the shortest route, though this did not reach statistical significance, $r(69) = -.19$,

p = .113. The correlation between the deviation above the shortest route and spatial working memory was also not significant, $r(68) = -.14$, $p = .239$.

When performing the task without the option to offload cognition via pen (i.e., in the no-pen and no-hand conditions), participants with better spatial working memory (Symmetry Span score) were expected to exhibit less deviation above the shortest route.

In the test blocks precluding pen use (i.e., no-pen for participants in the pen→no-pen and no-pen→pen groups, and no-hand for participants in the pen→no-hand and no-hand→pen groups), participants with greater spatial working memory had lower deviation above the shortest route, $r(134) = -.24$, $p = .005$. When examining test conditions separately, the correlation was still significant for the no-hand condition, $r(67) = -.31$, $p = .010$, but not for the no-pen condition, $r(65) = -.21$, $p = .089$.

Similarly, in the training blocks without pen use (i.e., no-pen for the no-pen→pen group, and no-hand for the no-hand→pen group), better spatial working memory was associated with smaller deviations from the shortest route, $r(64) = -.30$, $p = .016$. Again, when the training conditions were considered separately, the correlation remained significant for the no-hand condition, $r(31) = -.52$, $p = .002$, but not for the no-pen condition, $r(31) = -.12$, $p = .511$.

Overall, the results are generally consistent with our expectations, though some variations were observed across blocks. Indeed, better visual search abilities were associated with a greater use of offloading strategies in the pen condition, but this relationship reached statistical significance only in the training block. A greater use of offloading strategies was significantly related to a smaller deviation above the shortest route, though this effect was found only in the test block.

In the no-pen conditions, better spatial working memory was generally associated with smaller deviations from the optimal route. This relationship held both in the combined test conditions and specifically in the no-hand condition, for both the test and training blocks. The absence of significance in the no-pen condition may reflect insufficient statistical power or subtle differences in the cognitive demands between the two non-offloading conditions.

1. Study 2 – Consistency of the Offloading Strategies Devised by Participants Trained With Pen Use Between the Training and the Test-Block

To assess whether the offloading strategies devised by participants trained with pen use (pen→no-pen and pen→no-hand groups, N = 71) were consistently employed between the training and the training-consistent test blocks, and whether this consistency varied as a function of the presentation order of test blocks, we conducted a multiple regression analysis. In the model, we regressed the number of offloading strategies devised during the training block, the presentation order of test blocks (training-consistent → training-inconsistent *vs*. training-inconsistent → training-consistent), and their interaction term, on the number of offloading strategies used in the training-consistent test block. The analysis revealed that the number of offloading strategies used in the training block ($M = 3.37$, $SD = 0.90$) was a very strong predictor of the number of offloading strategies used in the training-consistent test block ($M = 3.44$, $SD = 0.84$), $\beta = .924$, $p < .001$, while neither the presentation order of test blocks (training-consistent → training-inconsistent, $M = 3.52$, $SD = 0.99$; training-inconsistent → training-consistent, $M = 3.34$, $SD = 0.64$), $\beta = -.221$, $p = .185$, nor the interaction term, $\beta = -.224$, $p = .073$, were significant predictors.

These findings indicate that, in the training-consistent test block, participants trained with pen use predominantly employed the same offloading strategies they had devised during the training. This may be compatible with the idea that interference from offloading strategies arises when transitioning to a more taxing, primarily working memory-based strategy for route planning, originating in the training phase (as suggested by this analysis) and continuing into the test phase (as indicated by the mediation analysis). However, these findings may also align with the notion of difficulty in abandoning an initial strategy (Luchins & Luchins, 1950), particularly when it is grounded in intuitive visual routines.

APPENDIX C

1. Study 3 - Correlations Between Individual Differences in Cognitive Abilities and Route Planning Task With/Without External Aids for Offloading Strategies

To investigate whether higher visual search abilities were associated with a smaller deviation above the shortest route in the presence of external aids for offloading strategies, and whether better verbal working memory was related to shorter solution-typing time in the absence of such aids, we conducted correlation analyses. In Study 3, there were five different groups (full offloading → no offloading, partial offloading → no offloading, no offloading → full offloading, no offloading → partial offloading, no offloading → no offloading), but only three conditions: full offloading (external aids for all four offloading strategies), partial offloading (external aids for two offloading strategies), and no offloading (no aids for offloading strategies). To increase statistical power, these analyses were performed by condition rather than by group.

In the full offloading condition, better visual search abilities (MFTC accuracy score) were significantly associated with a smaller deviation above the shortest route, $r(65) = -.34$, $p = .006$, consistent with H2. However, this pattern was not observed in the partial offloading condition, $r(64) = .03$, $p = .807$.

In the no offloading condition, better verbal working memory (Backward Digit Span) was associated with shorter solution-typing time, $r(161) = -.18$, $p = .025$.

The correlation analyses provided further insights into transfer effects. In line with our assumptions, participants with better visual search abilities demonstrated more accurate route planning in the full offloading condition, reinforcing the idea that a predominantly visual strategy is used for planning the routes when visual aids are externally available. However, this correlation did not hold in the partial offloading condition. Further analyses indicated that better verbal working memory was significantly associated with less time spent reporting the solution in the partial offloading condition ($r(65) = -.24$, $p = .047$), but not in the full offloading condition ($r(65) = -.21$, $p = .091$). This may suggest that, when only two strategies were externally aided, participants might have relied on a hybrid strategy. Furthermore, better verbal working memory was associated with shorter solution-typing times in the no offloading condition, consistent with our expectations, and underscoring the role of working memory in supporting performance when external aids are not available.

APPENDIX D

1. Study 4 - Analyses on Time Measures from Recall Tasks

The following tables reports regression analyses on times measures from recall tasks of Study 4. Like for analyses on accuracy measures, in the first block we entered group as predictor, in the second we included additional support use, in the third their interaction term and, in the fourth, we added recall task presentation order.

Table A.D.1

Regression Analysis on the Time Needed to Recall the Planned Route in the Verbal Task

Model	R	ΔR²	R²	Adjusted R²	F	df1	df2	p
1	0.11		0.01	0.00	1.20	1	91	.275
2	0.17		0.03	0.01	1.32	2	90	.273
1 - 2		0.02			1.42	1	90	.236
3	0.17		0.03	0.00	0.90	3	89	.446
2 - 3		0.00			0.08	1	89	.773
4	0.19		0.03	-0.01	0.79	4	88	.535
3 - 4		0.01			0.48	1	88	.488

None of the models tested in the regression analysis explained a significant amount of variance in the time needed to verbally recall the planned route.

Table A.D.2

Regression Analysis on the Time Needed to Recall the Planned Route in the Visual Task

Model	R	ΔR²	R²	Adjusted R²	F	df1	df2	p
1	.05		.00	-.01	0.26	1	91	.609
2	.19		.04	.01	1.64	2	90	.201
1 - 2		.03			3.00	1	90	.087
3	.28		.08	.05	2.53	3	89	.062
2 - 3		.04			4.19	1	89	.044
4	.30		.09	.05	2.15	4	88	.081
3 - 4		.01			1.02	1	88	.314

Although the inclusion of the interaction term between group and additional aid use in the third block explained significantly more variance than the block including group and additional aids use as individual predictors, none of the models tested in the regression analysis explained a significant amount of variance in the time needed to visually recall the planned route.

Table A.D.3

Regression Analysis on the Time Needed to Recall the Planned Route in the Paired Task

Model	R	ΔR^2	R^2	Adjusted R^2	F	df1	df2	p
1	.09		.01	.00	0.77	1	91	.382
2	.11		.01	-.01	0.51	2	90	.601
1 - 2		.00			0.26	1	90	.614
3	.18		.03	.00	0.95	3	89	.418
2 - 3		.02			1.83	1	89	.180
4	.18		.03	-.01	0.72	4	88	.580
3 - 4		.00			0.05	1	88	.822

None of the models tested in the regression analysis explained a significant amount of variance in the time needed to recall the planned route in the paired task.