Individual Differences in Working Memory Capacity Predict Benefits to Memory from Intention Offloading

SCHOLARONE™
Manuscripts

1

Individual Differences in Working Memory Capacity Predict Benefits to Memory

from Intention Offloading

Hunter Ball[1], Phil Peper[1], Durna Alakbarova[1], Gene Brewer[2], Sam J. Gilbert[3]

[1]University of Texas at Arlington

[2]Arizona State University

[3]University College London

Running Head: INTENTION OFFLOADING

2

Abstract

Research suggests that individuals with lower working memory have difficulty remembering to fulfill delayed intentions. The current study examined whether the ability to offload intentions onto the environment mitigated these deficits. Participants (N = 268) completed three versions of a delayed intention task with and without the use of reminders, along with three measures of working memory capacity. Results showed that individuals with higher working memory fulfilled more intentions when having to rely on their own memory, but this difference was eliminated when offloading was permitted. Individuals with lower working memory chose to offload more often, suggesting that they were less willing to engage in effortful maintenance of internal representations when given the option. Working memory was not associated with metacognitive confidence or optimal offloading decisions based on point value. These findings suggest offloading may help circumvent capacity limitations associated with maintaining and remembering delayed intentions.

Keywords: prospective memory; offloading; reminders; working memory; individual differences

Word count: 8,311

Individual Differences in Working Memory Capacity Predict Benefits to Memory from

Intention Offloading

Prospective memory refers to the ability to remember to perform delayed intentions at the appropriate moment in the future (e.g., take medication after dinner). Prospective memory differs from retrospective memory in that there is no explicit retrieval cue to query memory at the appropriate moment. Rather, with prospective memory the individual must self-initiate retrieval of the intention (Craik et al., 1986; Craik & McDowd, 1987), for example, upon noticing the medicine bottle on the counter while preparing dinner. This lack of environmental support often necessitates the use of demanding processes to maintain the intention in working memory (Ball & Brewer, 2018; Gilbert et al., 2020). This produces an unfortunate predicament: maintaining the intention in focal awareness could reduce processing resources available for performing ongoing activities (e.g., preparing dinner), but conserving processing resources poses the risk of forgetting the prospective memory intention. One solution to this problem is to *offload* demands onto the environment. For example, a neon sticky note can be posted near the medicine bottle, which in turn can stimulate intention retrieval in a more automatic fashion by reducing prospective memory processing demands (i.e., making it easier to notice the medicine bottle). (Gilbert et al., 2020; Risko & Gilbert, 2016; Scarampi & Gilbert, 2021). The purpose of the current study is to better understand each of these issues.

Intention Offloading and Biases

One way to reduce internal processing demands is to offload cognitive demands onto the external environment (Risko & Gilbert, 2016). Because typical laboratory prospective memory tasks are designed in such a way to prevent participants from any type of offloading, Gilbert and colleagues developed a task that permits intention offloading (e.g., Gilbert, 2015a; Gilbert et al.,

2020). In the delayed intention task, participants drag circles with letters or numbers to the

bottom of the screen in a specified order (e.g., 1, 2, 3…). Periodically a target circle appears in a

different color (e.g., orange) indicating it should later be dragged to a different location that

matches that color (e.g., upper orange location). During *internal* trials (i.e., no reminder),

participants must rely on their own memory to move the targets to the correct location at the

appropriate time. During *external* trials (i.e., reminders), participants are allowed to preemptively

move the target near the location to which it eventually needs to be dragged (i.e., offload

memory demands). Metacognitive performance predictions are also sometimes assessed by

having participants predict subsequent memory on internal trials. Consistent across all studies

using various versions of the task is that memory is better when offloading is permitted (Boldt &

Gilbert, 2019; Gilbert, 2015a, 2015b; Gilbert et al., 2020; Landsiedel & Gilbert, 2015; Sachdeva

& Gilbert, 2020; Scarampi & Gilbert, 2020, 2021). Additionally, participants are generally

underconfident in their internal ability (Boldt & Gilbert, 2019), which is often found in standard

prospective memory tasks as well (Meeks et al., 2007; Schnitzspahn et al., 2011; Susser &

Mulligan, 2019). Finally, the proportion of trials participants choose to offload on (i.e.,

offloading proportion) generally increases with increased task demands or reduced metacognitive

confidence (Gilbert, 2015a, 2015c; Gilbert et al., 2020; Sachdeva & Gilbert, 2020; Scarampi &

Gilbert, 2021).

While offloading can increase remembering or free attention resources for ongoing task

processing (Loft et al., 2011), using reminders can also be costly in terms of time and effort to

set up (Risko & Dunn, 2015; Risko & Gilbert, 2016). Optimal decision making should therefore

evaluate both the costs and benefits of offloading. Gilbert et al. (2020) argue that decisions to

offload might occur for at least two reasons: *metacognitive bias* and *effort minimization*.

Metacognition refers to the monitoring and control of one's own cognition during acquisition, retention, and retrieval of information (Nelson et al., 1990). Metacognition is central to many theories of how prospective memory intentions are realized, as it is used to inform strategy selection or to determine the appropriate attention allocation policy (Kuhlmann, 2019; Marsh et al., 2005; Penningroth & Scott, 2013; Rummel et al., 2019; Shelton et al., 2019). If beliefs about one's memory ability are low (Touron, 2015) or if demands on memory maintenance are perceived as intrinsically costly (Shenhav et al., 2017), one may opt to use a strategy that avoids reliance on memory. Accordingly, the metacognitive bias view suggests that underconfidence in one's own internal memory leads to compensatory offloading to ensure intention completion (Dunn & Risko, 2016; Gilbert et al., 2020; Risko & Dunn, 2015), whereas the effort minimization view suggests that participants may offload to minimize the amount of effort it takes to complete the task (Kool et al., 2010; Sachdeva & Gilbert, 2020; Shenhav et al., 2017). In both cases, an overreliance on offloading can result in unnecessary usage of costly resources to set up reminders, referred to as "reminder bias".

To explore optimal decision making, Gilbert et al. (2020) modified the standard delayed intention task to include both forced and choice trials. In the modified task participants are required to either rely on their own memory (forced internal) or to offload (forced external) on forced trials, but are free to choose between the two (choice internal or external) on choice trials. Critically, choosing to offload on choice trials results in fewer points for each successfully moved target (e.g., ranging from 1-9 points) than relying on one's own memory (always worth 10 points). Because forced internal memory is around 50% and external memory is near 100%, this allows for an easy determination of optimal values to offload. For example, if a participant can remember an average of 5 out of 10 targets on forced internal trials, choosing to offload at a

value of 4 would be suboptimal and is indicative of a *reminder bias*. Across three experiments, Gilbert and colleagues found that participants were biased to use reminders as indicated by the choice to offload at point values lower than what was optimal based on their actual internal memory performance. Underconfidence was also associated with a greater reminder bias. Interestingly, the reminder bias was eliminated by metacognitive advice on which strategy to use on each choice trial and was reduced (but not eliminated) by increasing internal memory confidence following easy practice and positive performance feedback. These findings are generally consistent with the metacognitive bias view, whereby inaccurate assessments of one's own memory ability can lead to an overreliance on reminders for intentions that may otherwise be completed more efficiently without offloading.

The finding that positive feedback and easy practice did not eliminate the reminder bias suggests that other factors may have also contributed to overreliance on offloading. Sachdeva and Gilbert (2020) replicated the procedure and compared it to a condition in which participants received payment contingent on the total number of points scored. While participants in both conditions were equally confident, those with a financial incentive showed a significant reduction in reminder bias. This suggests that the preference to avoid cognitive effort may also underlie the reminder bias, which can be reduced with the appropriate incentive to rely on internal memory. While these findings suggest that both metacognitive bias and effort minimization underlie offloading decisions, less is known about the specific cognitive processes that give rise to these decisions.

Individual Differences in Working Memory

Working memory, broadly defined, refers to the attention and memory control processes needed to maintain goal-relevant information in focal awareness and to retrieve from long-term

memory information displaced due to distraction (Kane et al., 2001; Kane & Engle, 2003; Kane et al., 2004; Unsworth, Brewer, et al., 2012; Unsworth & Engle, 2007). Considerable research has shown that individuals with higher working memory outperform those with lower working memory on cognitive tasks that place high demands on internal processes, but these effects are mitigated with sufficient environmental support. For example, working memory differences arise during anti-saccade tasks where salient features conflict with task goals, but these differences do not occur on prosaccade tasks where the environmental cues facilitate task goals (Unsworth et al., 2004). Similarly, differences are evidenced in free recall tasks that place high demands on self-initiated retrieval, but are attenuated when participants are given category cues to facilitate retrieval (Unsworth, Spillers, et al., 2012). The findings suggest that individuals with lower working memory ability have greater difficulty in maintaining internal representations when attention and memory demands are sufficiently high (or environmental support is low).

Working memory has also been shown to predict performance on standard prospective memory tasks that place high demands on attention or memory (Arnold et al., 2015; Ball & Brewer, 2018; Ball et al., 2013; Brewer et al., 2010; Knight et al., 2012; McDaniel et al., 2013; Reynolds et al., 2009; Smith & Bayen, 2005; Unsworth, Brewer, et al., 2012). It is argued that the reason working memory is predictive of performance is because similar controlled processes are often needed to complete both working memory tasks and prospective memory tasks (Ball et al., 2019; Marsh & Hicks, 1998). For example, Brewer et al. (2010) found that higher working memory participants outperformed lower working memory participants on nonfocal tasks that required attentionally demanding preparatory monitoring processes, but not on focal tasks where intention retrieval could occur relatively automatically (see also Arnold et al., 2015; Ball & Brewer, 2018; Ball et al., 2013; Unsworth, Spillers, et al., 2012). Additionally, Ball et al. (2018)

found that higher working memory participants had fewer errors of commission or omission than

lower working memory participants when difficult memory search was needed to determine

whether an intention had previously been fulfilled (see also Ball et al., 2013). These findings

suggest that working memory is critical for maintaining and retrieving internal representations

associated with prospective memory tasks. Consequently, individuals with poorer working

memory stand to benefit more from offloading onto the environment. However, just because

offloading *can* help, this does not mean that individuals with lower working memory will

necessarily engage in appropriate compensatory strategies to aid memory (Morrison &

Richmond, 2020).

No studies to date have directly examined the role of working memory in benefits from

offloading or decisions to offload during the delayed intention task. Morrison and Richmond

(2020) had participants perform a short-term memory serial recall task with letters of various set

sizes (e.g., ranging from 2-10 letters) during forced internal (no reminder) and choice blocks.

Although offloading improved short-term memory performance, particularly under high loads

(e.g., set sizes of 6, 8 and 10), this did not differ as a function of working memory. Moreover,

working memory was not related to the proportion of trials offloaded (but see Risko & Dunn,

2015). Scarampi and Gilbert (2020) showed nearly an identical pattern of results comparing

younger and older adults in a delayed intention task: offloading improved memory for delayed

intentions (particularly under high load), offloading benefits did not differ as a function of age,

and age was not associated with offloading proportion. Notably, older adults were overconfident

in their memory ability, suggesting that they may not have been aware of the potential utility of

offloading to reduce memory declines. It is possible that similar lack of metacognitive awareness

resulted in the null relation between working memory and short-term memory offloading

proportion seen by Morrison and Richmond (2020). It remains an open question whether

working memory differences might be evidenced in the delayed intention task and whether

metacognitive biases and/or effort minimization might underlie these differences.

Current Study

The purpose of the current study was to examine how offloading influences memory for

delayed intentions in individuals differing in working memory ability and how cognitive ability

influences decisions to offload. We examined memory for delayed intentions using a large-scale

individual differences design with multiple performance indicators for each construct.

Participants completed three variants of the delayed intention task ('ABC', '123', and '321'; see

below for details), similar to Gilbert et al. (2020), that included both forced and choice trials with

point values. Additionally, participants completed three computerized complex span tasks that

are commonly used to assess working memory ability (Operation, Reading, and Symmetry

Span). Using multiple performance indicators allows for the use of latent variable modeling,

which is a useful approach because it controls for measurement error while testing different

theoretical predictions of the relation between working memory ability, intention memory, and

offloading decisions.

Regarding intention memory, we anticipated that without offloading (i.e., forced internal

trials) individuals with higher working memory would successfully fulfill more intentions than

those with lower working memory (Arnold et al., 2015; Ball & Brewer, 2018; Ball et al., 2013;

Brewer et al., 2010; Knight et al., 2012; McDaniel et al., 2013; Reynolds et al., 2009; Smith &

Bayen, 2005; Unsworth, Brewer, et al., 2012). Critically, because offloading can be used to

circumvent capacity limitations associated with maintaining and remembering the intention,

working memory differences should be reduced or eliminated when offloading is required. This

should especially be the case in the current study, as Gilbert et al. (2020) showed that forced

offloading in this paradigm produced external memory rates near 100%. Of course, it is possible

that participants may not effectively use reminders, meaning that working memory differences

could still arise (Morrison & Richmond, 2020; Scarampi & Gilbert, 2021).

Regarding offloading decisions during choice trials, we expected that participants would

show a reminder bias and for this to be greater for those less confident in their internal memory

ability (Gilbert et al., 2020). We also hypothesized that individuals with lower working memory

would offload more often independent of point values (i.e., offloading proportion) and would be

less optimal in their offloading decisions when trying to earn points (i.e., reminder bias). A

preference towards externalizing cognition for low working memory participants could be driven

by an underconfidence in their own memory ability or desire to minimize the amount of effort

expended to complete the task. The former would be evidenced by a positive correlation between

working memory and underconfidence (i.e., lower working memory is associated with greater

underconfidence), which in turn, results in a greater reminder bias. The latter would be

evidenced by finding that working memory is associated with offloading decisions but unrelated

to confidence.

Methods

All research reported herein was conducted using appropriate ethical guidelines and was

approved by the Institutional Review Board at the University of Texas at Arlington. We report

how we determined our sample size, all data exclusions, and all manipulations.

Participants and Design

The current study was conducted in the context of a larger cognitive battery for a separate

experiment. The study consisted of two sessions scheduled one week apart, each lasting

1

approximately two hours. A desired sample size of at least 250 over the course of two semesters

was chosen based on recommendations that 250 participants are needed to detect stable and

reliable correlations (Schönbrodt & Perugini, 2013). We applied a stopping rule beyond 250

participants that coincided with the end of the second semester. Over the course of two

semesters, a total of 310 undergraduates from the University of Texas Arlington enrolled in the

study to receive participation credit towards course requirements, but only 279 participants

completed both days. After participant exclusions (described below), the final sample consisted

of 268 participants.

Materials and Procedure

Delayed intention tasks

The materials and stimuli for the three delayed intention tasks were adapted from Gilbert

et al. (2020). The only difference across the three tasks were the actual stimuli. In the letters task,

participants dragged circles containing letters in alphabetical order (A, B, C…). In the numbers

ascending task, they dragged circles containing numbers in ascending order (1, 2, 3…). In the

numbers descending task, they dragged circles containing numbers in descending order (17, 16,

15…). Below, we describe only the letters task in detail, as all three tasks followed an identical

procedure.
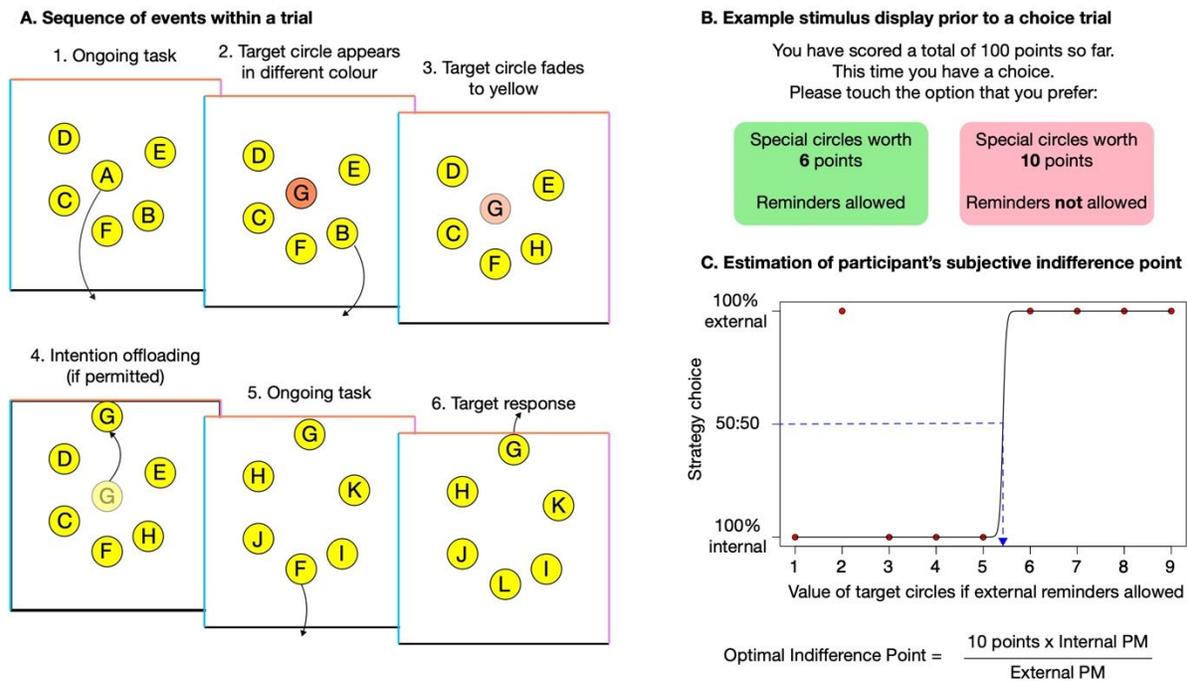
Figure 1. Example of the letter version of the delayed intention task.

As can be seen in Figure 1, during the ABC (letters) task six yellow circles were presented within a square on the computer. Each circle contained a letter from the alphabet and participants were to drag the circles sequentially (in alphabetical order) to the bottom of the square. Each time a circle was dragged to the bottom of the square, a new circle appeared in its original location, continuing the alphabetical sequence. This continued until 17 circles were dragged out of the square (i.e., letters A-Q). Occasionally, new circles (i.e., targets) initially appeared in blue, orange, or pink, rather than yellow, which corresponded with the left, top, and right side of the square, respectively. Two seconds after appearing on the screen, the color faded to yellow so that they matched the other circles. When a target appeared (e.g., in blue), this represented an instruction that it should eventually be dragged to its corresponding side of the square (e.g., left) when it was reached in the alphabetical sequence. For example, a participant

first drags A to the bottom of the screen where it disappears. A blue G appears in its place,

fading to yellow after 2 seconds. Meanwhile, the participant drags circles B-F to the bottom of

the screen, before dragging G to the left. Importantly, targets can be remembered in two different

ways. Participants can rely on their own internal representation of where it should eventually be

dragged (i.e., no reminder). Alternatively, participants can set an external reminder as soon as it

appears by moving it near the location (e.g., left side) where it eventually needs to be dragged.

An everyday analogy would be leaving an object by your front door so that you remember to

take it with you next time you leave the house. This was manipulated in the current study

(described below).

One trial consisted of a 17-letter alphabetical sequence (A-Q). Within this sequence,

a total of 6 target circles appeared, with the 6 letters randomly allocated from G to Q. This means

that participants had to remember multiple simultaneous intentions. The 6 target circles were

randomly allocated to the left, top, and right positions of the square. Feedback was provided by

the circle changing color before disappearing if dragged to the correct location (green) or

incorrect location (red). All circles correctly dragged to the bottom of the box turned purple

before disappearing. Note that the number ascending and number descending tasks were

performed in the exact same way, except it consisted of the numbers 1-17 (ascending) or 17-1

(descending) rather than letters. For a demonstration of the letter task, please visit:

http://samgilbert.net/optimalDemo/start.html.

For each task (i.e., letters, numbers ascending, or numbers descending), participants

performed a total of 11 experimental trials following a brief practice session, where each trial

consisted of a full set of 17 circles including 6 targets. Participants were forced to use an internal

(unaided memory) strategy for 3 trials or an external (reminder) strategy for 3 trials; during the

other 7 trials participants were free to choose between internal and external strategies. To force

an internal strategy, circles were fixed in position on the screen (other than the current one that

needed to be dragged in sequence) so that target circles could not be moved when they first

appeared. To force an external strategy, when a target circle appeared the task could only be

continued after the participant moved it within the square. Prior to beginning a forced internal or

external trial, participants were informed which strategy they had to use. Participants were told

that they scored points every time they dragged one of the target circles to the instructed

location. On trials where they were forced to use an internal or external strategy, they scored 10

points for each correct target response. These conditions occurred on trials 2, 4, 6, 8, and 10,

alternating between internal and external (the order of which was counterbalanced across

participants). On the remaining seven trials, participants were given a free choice (see Figure 1,

panel B for an example). For the choice trials, they could choose to use an internal strategy for

the upcoming trial, in which case they scored 10 points per correct target response but were

prevented from setting external reminders. Alternatively, they could choose to set reminders in

the upcoming trial, in which case they were offered a lower number of points – randomly

ranging from 2 to 8 – for each correct target response. After each trial, participants were told the

total number of points that they had scored in the experiment so far. They were told to try to

score as many points as possible, and that on choice trials they should choose whichever strategy

they believed would allow them to score more points.

The actual order of the three task versions began with numbers ascending (123), followed

by letters (ABC), and finally, numbers descending (321). After receiving the ongoing task

instructions for the first (numbers ascending) task, participants performed a practice trial with 8

circles where they moved the numbers, in order, to the bottom of the screen. Participants were

then given the instructions, followed by an 8-circle practice where the last circle was a target.

They then performed a full 17-circle practice phase with 6 targets, without the use of reminders.

Following this practice, participants were asked to predict what percentage of target circles (from

0-100%) they thought they would remember to drag to the appropriate side of the square during

the actual task. Finally, participants were given instructions on how to set reminders and

performed a full 17-circle practice phase where reminder usage was required. They then

completed the 11 trials as described above. Upon completion of this task, participants

immediately started the next task version (i.e., ABC). Because participants were already familiar

with the task structure at this point, they only completed the full 17-circle practice phase without

reminders. Following this practice, participants made predictions (0-100%) on how they thought

they would do in the actual task without reminders. They then completed the 11-trial procedure

for the letters task. Finally, the same procedure was repeated for the last (i.e., 321) task.

Dependent Variables

Internal Memory. Internal memory was calculated as the proportion of target circles

correctly dragged to their instructed locations on forced internal trials.

External Memory. External memory was calculated as the proportion of target circles

correctly dragged to their instructed locations on forced external trials.

Offloading proportion: Offloading proportion was calculated by dividing the total

number of times participants opted to use a reminder by the total number of choice trials (i.e.,

seven).

Optimal indifference point. The expected score on forced internal trials is 10 x actual

accuracy on these trials (since each target was worth 10 points). The *optimal* indifference point is

the target value that would lead participants to achieve the same score if they are allowed to use

reminders (i.e. [optimal indifference point]*[accuracy on forced external trials]. Therefore, [optimal indifference point]*[accuracy on the forced external trials] = [10]*[accuracy on the forced internal trials].

Actual indifference point. To calculate the *actual* indifference point (i.e., the value at which participants were equally likely to choose an internal or an external strategy) we calculated the likelihood of choosing an external vs internal strategy across the full range of external target values from 2-8. We then fit a sigmoid function to these data using the R package 'quickpsy', bounded to the range 2-8 and otherwise using default parameters. This allowed us to calculate the value associated with a 50% probability of choosing either strategy, according to this function.

Reminder bias: Reminder bias was calculated as the difference between the optimal and actual indifference scores. A score of zero means they are unbiased, a positive score means they are biased to rely on an external strategy, and a negative score indicates they are biased to rely on an internal strategy.

Metacognitive bias. Metacognitive bias was calculated as the difference between predictions (i.e., global confidence) and actual performance on forced internal trials. A score of zero reflects they are unbiased, a negative score means they are underconfident, and a positive score means they are overconfident.

Working Memory Tasks

For each working memory measure, participants first engaged in a three-part practice, where they first practiced the storage component of the task alone, then they practiced the processing component of the task alone, and finally, they completed the processing component followed by the storage component. Each trial in the actual working memory tasks was presented

1

for a length of time equal to 2.5 standard deviations above the mean for response times in that

task's processing-only practice trials. Abbreviated versions of each task were used, such that

during the actual task there were two trials of each list length (Oswald et al., 2015). The

dependent variable for all tasks was the proportion of correct items in the correct serial position.

Reading Span. Participants were required to read sentences while trying to remember a

set of unrelated letters. For this task, participants read a sentence and determined whether the

sentence made sense or not (e.g., "The prosecutor's dish was lost because it was not based on

fact?"). Half of the sentences made sense while the other half did not. Nonsense sentences were

made by simply changing one word (e.g., "dish" from "case") from an otherwise normal

sentence. Participants were required to read the sentence and to indicate whether it made

sense. After participants gave their response, they were presented with a letter for 1000 ms. At

recall, the letters from the current set were recalled in the correct order by clicking via the mouse

on the appropriate letters displayed on the computer screen. There were two trials of each list-

length with the list-length ranging from 3–7 letters.

Operation Span. Participants solved a series of math operations while trying to

remember a set of unrelated letters (F, H, J, K, L, N, P, Q, R, S, T, Y). Participants were required

to solve a math operation, and after solving the operation they were presented with a letter for 1

s. Immediately after the letter was presented the next operation was presented. At recall, letters

from the current set were recalled in the correct order by clicking via the mouse on the

appropriate letters displayed on the computer screen. Participants received three sets (of list-

length two) of practice. There were two trials of each list-length with the list-length ranging

from 3–7 letters.

1

Symmetry Span. In this task, participants were required to recall sequences of red squares within a matrix while performing a symmetry-judgment task. In the symmetry-judgment task, participants were shown an 8 x 8 matrix with some squares filled in black. Participants decided whether the design was symmetrical across its vertical axis. The pattern was symmetrical half of the time. Immediately after determining whether the pattern was symmetrical, participants were presented with a 4 x 4 matrix with one of the cells filled in red for 650 ms. At recall, participants recalled the sequence of red-square locations in the preceding displays, in the order they appeared, by clicking via the mouse on the cells of an empty matrix displayed on the computer screen. There were two trials of each list-length, with the list-length ranging from 2-5 squares.

Data Analytic Approach and Participant Exclusions

Data analysis. Confirmatory factor analysis (CFA) was used to examine the relations among the various delayed intention tasks and working memory. CFA is a robust analytic technique that reduces spurious relations among measures based on task-specific variance or measurement error. In this approach, a theoretically derived model is specified and the corresponding hypothetical variance-covariance matrix is compared with the true variance-covariance matrix for the constructs of interest (Kline, 2015). How well the specified model reproduces the observed data can be assessed using a chi-square statistic and goodness-of-fit indices (described later). Missing data was imputed using maximum likelihood estimation.

Participant exclusions. A total of 279 participants completed both days of the study. Participants were subsequently excluded from analyses for the following reasons: failing to complete any of the delayed intention tasks ($n = 1$) or failing to maintain an average of 70%

accuracy across the laboratory tasks on external (reminder) trials ($n = 10$)[1]. The accuracy

exclusion was based on criteria specified by Gilbert et al. (2020), as such low performance likely

reflects a failure to follow instructions. The final sample, therefore, consisted of 268 participants.

Results

Task Level Results

Descriptive statistics for all variables are listed in Table 1. All measures had acceptable

values of skew and kurtosis (skew $< |3|$ and kurtosis $< |8|$; Kline, 2011). Correlations for the

primary measures of interest can be found in Table 2, and scatter plots are displayed in the

supplemental material.

Intention Memory. Across all three tasks, external memory accuracy was higher than

internal accuracy, indicating that participants benefited from offloading (IOT_{123}: $F(1,267) =$

1338.92, $p < .001$, $\eta_p^2 = .834$; IOT_{ABC}: $F(1,267) = 1680.49$, $p < .001$, $\eta_p^2 = .863$; IOT_{321}:

$F(1,267) = 1256.19$, $p < .001$, $\eta_p^2 = .825$).

Reminder Bias. Across all three task, actual indifference points were lower than optimal

indifference points, indicating that participants were biased to offload at lower point values than

they should have based on their actual memory ability (*reminder bias*: IOT_{123}: $F(1,267) = 66.02$

$p < .001$, $\eta_p^2 = .198$; IOT_{ABC}: $F(1,267) = 83.34$, $p < .001$, $\eta_p^2 = .238$; IOT_{321}: $F(1,267) = 95.08$, p

$< .001$, $\eta_p^2 = .263$).

---

[1]Prior research also excluded participants with negative correlations between target value and the likelihood of choosing reminders, as this reflects a random or counter-rational choice strategy (Gilbert et al., 2020; Sachdeva & Gilbert, 2020). In the current study, averaged across all three tasks there were 25 participants that had a negative correlation between the two variables. Because the focus of the current study is on individual differences in offloading decisions, we opted to retain these participants. Note that excluding these participants resulted in an identical pattern of results as reported below. There was also one participant that had equivalent average performance between internal and external trials that we did not exclude because this made no influence on the results.

Metacognitive Bias. Across two of the three tasks, predictions of internal memory were lower than actual internal memory, indicating that participants were underconfident in their memory ability (*metacognitive bias*: IOT_{123}: $F(1,267) = 1.38$, $p = .241$, $\eta_p^2 = .005$; IOT_{ABC}: $F(1,267) = 14.29$, $p < .001$, $\eta_p^2 = .051$; IOT_{321}: $F(1,267) = 24.45$, $p < .001$, $\eta_p^2 = .084$).

Table 1. *Descriptive Statistics for Intention Offloading and Working Memory Tasks*

Measure	Mean	SD	Min	Max	Skew	Kurtosis	Reliability
Internal - 123	0.57	0.15	0.17	1.00	0.14	-0.21	0.63
Internal - ABC	0.56	0.16	0.17	1.00	0.39	-0.04	0.73
Internal - 321	0.57	0.19	0.11	1.00	0.29	-0.32	0.77
External - 123	0.92	0.11	0.44	1.00	-1.79	3.15	0.77
External - ABC	0.96	0.06	0.72	1.00	-1.81	3.41	0.74
External - 321	0.97	0.06	0.67	1.00	-2.58	7.91	0.73
Proportion - 123	0.51	0.28	0.00	1.00	-0.01	-0.80	0.70
Proportion - ABC	0.60	0.31	0.00	1.00	-0.36	-1.00	0.79
Proportion - 321	0.63	0.33	0.00	1.00	-0.50	-0.91	0.83
OIP - 123	6.13	1.35	2.14	8.00	-0.36	-0.61	-
OIP - ABC	5.73	1.44	1.76	8.00	-0.11	-0.64	-
OIP - 321	5.71	1.59	2.00	8.00	-0.18	-0.69	-
AIP - 123	4.95	2.31	2.00	8.00	0.02	-1.51	-
AIP - ABC	4.31	2.34	2.00	8.00	0.46	-1.39	-
AIP - 321	4.18	2.35	2.00	8.00	0.59	-1.26	-
Prediction - 123	0.55	0.21	0.04	1.00	-0.14	-0.69	-
Prediction - ABC	0.50	0.25	0.00	1.00	0.08	-0.69	-
Prediction - 321	0.47	0.27	0.00	1.00	0.17	-0.90	-
Reminder Bias - 123	1.20	2.38	-5.06	6.00	-0.19	-0.60	-
Reminder Bias - ABC	1.43	2.55	-5.69	6.00	-0.46	-0.40	-
Reminder Bias - 321	1.55	2.56	-6.00	6.33	-0.42	-0.26	-
Metacognitive Bias - 123	-0.02	0.23	-0.64	0.56	-0.18	-0.32	-
Metacognitive Bias - ABC	-0.06	0.28	-0.67	0.67	0.10	-0.47	-
Metacognitive Bias - 321	-0.10	0.32	-0.89	0.78	0.25	-0.26	-
Ospan	0.71	0.22	0.00	1.00	-0.96	0.36	0.86
Rspan	0.67	0.17	0.02	1.00	-0.63	0.86	0.71
Sspan	0.63	0.20	0.07	1.00	-0.42	-0.31	0.70

Note. 123 = number ascending task; ABC = alphabetical task; 321 = descending task; OIP = optimal indifference point; AIP = actual indifference point; Ospan = operation span task; Rspan = reading span task; Sspan = symmetry span task. Reliability reflects Cronbach's Alpha for all measures where there were multiple assessments of performance.

Table 2. *Correlations Between Intention Offloading Measures and Working Memory.*

Meaure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Internal - 123	1																	
2. Internal - ABC	.48**	1																
3. Internal - 321	.40**	.47**	1															
4. External - 123	.31**	.25**	.18**	1.00														
5. External - ABC	.23**	.22**	0.11	.25**	1													
6. External - 321	0.10	.15*	.18**	.24**	.22**	1												
7. Proportion - 123	-.19**	-.19**	-.20**	.21**	0.06	0.05	1											
8. Proportion - ABC	-0.09	-.17**	-.15*	.15*	.15*	0.09	.71**	1										
9. Proportion - 321	-0.06	-0.12	-.21**	0.07	0.10	0.02	.58**	.75**	1									
10. Reminder Bias - 123	.36**	0.01	0.00	.13*	.13*	0.00	.73**	.53**	.44**	1								
11. Reminder Bias - ABC	.18**	.43**	.13*	.25**	.16**	.14*	.48**	.75**	.53**	.44**	1							
12. Reminder Bias - 321	.22**	.20**	.42**	.15*	.12*	0.04	.39**	.54**	.74**	.41**	.55**	1						
13. Meta Bias - 123	-.46**	-.13*	-.16**	-.15*	-0.01	-.14*	0.04	0.05	0.09	-.22**	-0.03	-0.02	1					
14. Meta Bias - ABC	-.14*	-.47**	-.21**	-0.06	0.03	-.15*	.18**	.16**	.20**	0.11	-.17**	0.03	.35**	1				
15. Meta Bias - 321	-0.09	-.22**	-.51**	0.00	0.11	-.14*	.20**	.18**	.19**	.13*	-0.01	-.18**	.32**	.60**	1			
16. Ospan	.19**	.21**	0.10	-0.05	0.12	0.00	-.12*	-.15*	-0.10	0.01	-0.01	-0.03	0.03	-0.04	0.00	1		
17. Rspan	.15*	.18**	.12*	-0.07	0.10	0.05	-0.10	-.14*	-0.08	0.01	-0.01	-0.03	0.00	-0.02	0.01	.61**	1	
18. Sspan	.32**	.22**	.26**	0.09	.14*	.14*	-.20**	-.25**	-.18**	0.00	-0.05	-0.03	-0.10	-0.07	-.13*	.31**	.37**	1

Note. 123 = number ascending delayed intention task; ABC = alphabetical delayed intention task; 321 = descending delayed intention offloading task; Ospan = operation span task; Rspan = reading span task; Sspan = symmetry span task. ** = $p < .01$; * = $p < .05$.

Intention Memory as a Function of Reminders

The first set of analyses examined whether offloading improved intention memory and whether this differed as a function of working memory ability. To determine the factor structure of the data, we used CFA to specify two theoretically plausible models. In the two-factor model, performance from both internal and external memory trials loaded onto a single factor alone and working memory loaded onto a separate factor. This model is a "general" model, as it tests the hypothesis that the processes underlying intention memory are largely invariant across offloading conditions. A three-factor "offloading" model was also specified in which the internal and external memory performance loaded onto separate factors. This model tests the hypothesis that different mechanisms may underlie intention memory depending on whether offloading is possible. The top row of Table 3 displays the fits of each model.

Table 3. *Model Fits for Confirmatory Factor Analysis on Intention Memory, Offloading Proportion, and Biases.*

DV	Model	df	χ^2	p	CFI	NNFI	SRMR	RMSEA [90% CI]
Intention Memory	General PM	26	67.95	< .01	0.90	0.86	0.07	.08 [.06, .10]
	Internal vs. External	**24**	**46.67**	**< 0.01**	**0.95**	**0.92**	**0.06**	**.06 [.03, .09]**
Offloading Proportion	**General Proportion**	**8**	**10.17**	**0.25**	**0.97**	**0.99**	**0.05**	**.03 [.00, .08]**
Reminder and Metacognitive Biases	General Bias	26	275.89	< .01	0.57	0.40	0.12	.19 [.17, .21]
	Reminder vs. Metacognitive	24	117.61	< .01	0.84	0.76	0.06	.12 [.10, .14]
	Reminder vs. Metacognitive (Correlated Residual)	**21**	**17.39**	**0.69**	**>0.99**	**>0.99**	**0.03**	**<.001 [.00, .04]**

Note. Model in **bold** is best fitting model. df = degrees of freedom, χ^2 = chi squared; p = p value; CFI = comparative fit index; NNFI = non-normed fit index; ; SRMR = standardized root mean square residual; RMSEA = root mean square error. The model in **bold** is best fitting.

In the factor analytic approach, the χ^2 statistic reflects how well the specified model reproduces the variance-covariance structure of the observed data. In this case, a significant χ^2

test (p < .05) is undesirable because it means that the theoretical model does not accurately

reflect the observed structure. However, because large samples can often produce a significant χ^2

value despite being a good model fit, other goodness-of-fit indices are reported. CFI and NFI

values greater than .90, and SRMR and RMSEA values less than .08, are indicative of acceptable

fit (Kline, 2015)[2]. As can be seen in Table 3, the three-factor model provided a good fit to the

data. The fit of the two models were directly compared using a χ^2 difference test. In this case, a

significant difference between the χ^2 values indicates that three-factor "offloading" model

provided a significantly better fit than the two-factor "general" model, $\Delta\chi^2 (2) = 21.28$, p < .001.

Figure 2 displays the correlations across the different factors for the best fitting three-

factor model. Significant paths (p < .05) are indicated by solid lines and nonsignificant paths are

dashed. Critically, higher working memory was associated with better internal memory, but

working memory was not associated with external memory. A Wald Test confirmed that the

standardized path coefficient between working memory and intention memory was significantly

greater for the internal factor than the external factor, $\chi^2 (1) = 11.29$, $p < .001$. These findings

suggest offloading attenuated memory differences between individuals varying in working

memory ability. It should be noted, however, that the external memory was quite high (95%),

which may artificially attenuate the correlation with working memory.

---

[2] The root mean square error of approximation (RMSEA) and the standardized root mean square residual
(SRMR) both reflect the average squared deviation between the observed and reproduced covariances. In
addition, the non-normed fit index (NNFI) and the comparative fit index (CFI), both of which compare the fit
of the specified model to a baseline null model. NNFI, and CFI values greater than .90 and RMSEA and
SRMR values less than .08 are indicative of acceptable fit.
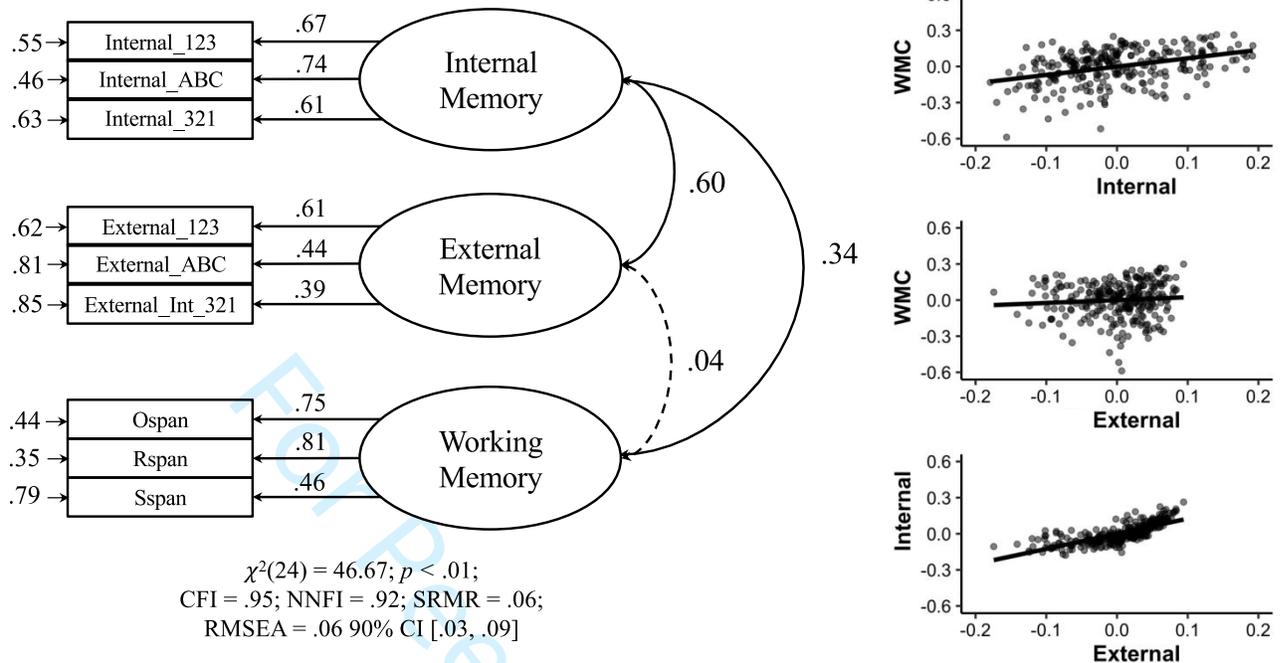
Figure 2. Best-fitting three factor confirmatory factor analysis examining the relation between internal memory, external memory, and working memory (left) and scatter plots of latent correlations (right). Solid lines in factor analysis indicate significant paths at $p < .05$, whereas dashed lines reflect non-significant paths.

Intention Offloading Proportion

The previous analysis indicated that working memory was associated with internal but not external memory. To determine whether this was due in part to low ability participants choosing to offload more when given the opportunity, we examined the relation between working memory and offloading proportion. Offloading proportion reflects the proportion of times (out of 7) participants opted to use reminders during the choice trials, independent of point values. To examine this, we specified a two-factor model where offloading proportion loaded onto one factor and working memory loaded onto another. As can be seen in the middle row of Table 3, this model provided an excellent fit to the data. As can be seen in Figure 3, individuals with lower working memory chose to offload more than did high working memory individuals.
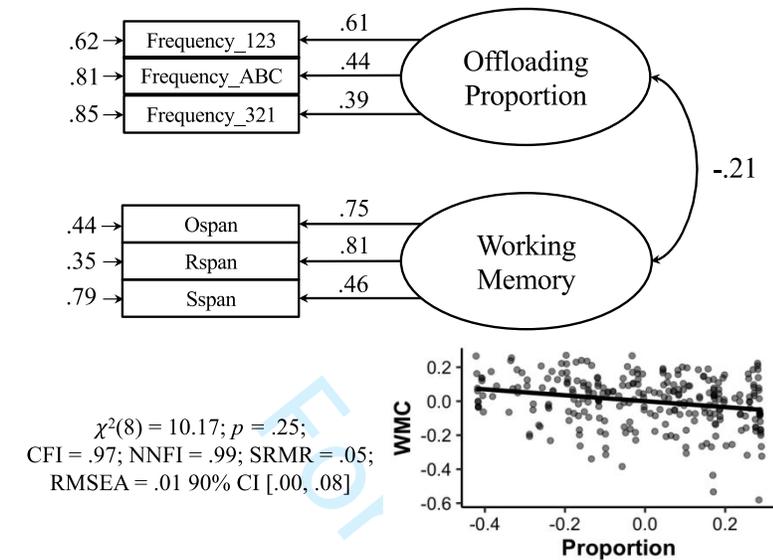
Figure 3. Confirmatory factor analysis examining the relation between offloading proportion and working memory (top) and scatter plots of latent correlations (bottom). Solid lines in factor analysis indicate significant paths at $p < .05$, whereas dashed lines reflect non-significant paths.

Optimal Offloading Decisions

The previous analysis indicated that low working memory participants offloaded more often. However, this proportion was calculated independent of point values. We were also interested in whether participants strategically used point values to offload optimally, whether this differed as a function of metacognitive confidence, and whether working memory contributed to these relations. Reminder bias reflects the difference between optimal and actual indifference points (with a positive score reflecting a bias to rely on an external strategy), whereas metacognitive bias reflects the difference between predictions and actual performance on internal trials (with a negative score reflecting greater underconfidence). Based on previous work, we would anticipate a negative relation between the two bias measures, meaning greater underconfidence leads to a bias to offload.

To examine the role of working memory ability in optimal offloading decisions, we specified a two-factor model where all delayed intention tasks loaded onto a single factor alone

and working memory loaded onto a separate factor. This model tests the hypothesis that the processes underlying metacognitive bias and offloading bias are not distinguishable. We compared this to a three-factor model where the delayed intention tasks loaded onto bias-specific factors (i.e., separate reminder bias and metacognitive bias factors). This model tests the hypothesis that different mechanisms underlie decisions to offload and confidence. Finally, we specified a three-factor model that allowed task-specific residuals to correlate across latent factors (e.g., ABC reminder bias and ABC metacognitive bias residuals were allowed to correlate). This model was created post-hoc based on the relatively poor fit of the initial three-factor model (Table 2) and examination of the task-level correlations. As can be seen in Table 1, reminder biases and metacognitive biases were significantly negatively correlated *within* a task (e.g., reminder bias on the 123 task was negatively correlated with the metacognitive bias on the 123 task), whereas the relations *across* tasks were either uncorrelated or even positively correlated (e.g., 123 reminder bias and ABC metacognitive bias).

As can be seen in bottom row of Table 3, the three-factor model with correlated residuals provided an excellent fit to the data. This model also provided a significantly better fit than the three-factor model with uncorrelated residuals, $\Delta\chi^2 (3) = 17.39$, $p < .001$. This suggests clear task-specificity in the association between reminder and metacognitive biases. Critically, as can be seen in Figure 4, none of the latent factors were significantly correlated with one another. This is inconsistent with previous research demonstrating that greater metacognitive bias leads to more reliance on external sources, although the effect is in that direction. In any manner, cognitive ability is clearly not related to either bias.

Running Head: INTENTION OFFLOADING
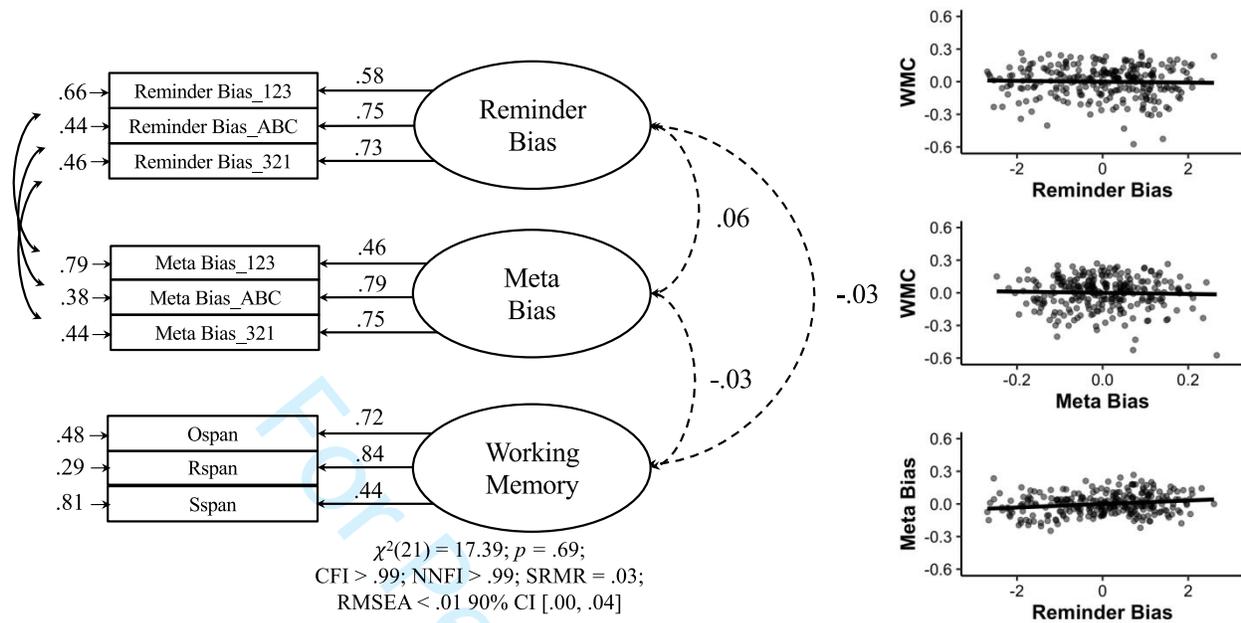
Figure 4. Best-fitting three factor confirmatory factor analysis model examining the relation between reminder bias, metacognitive bias, and working memory (left) and scatter plots of latent correlations (right). Solid lines in factor analysis indicate significant paths at $p < .05$, whereas dashed lines reflect non-significant paths.

General Discussion

The purpose of the current study was to better understand the mechanisms underlying intention offloading and who might benefit most from offloading. Participants performed multiple delayed intention tasks with and without the use of reminders. Consistent with prior research, reminders drastically improved intention fulfillment (Gilbert, 2015a, 2015b, 2015c; Gilbert et al., 2020; Haines et al., 2020; Ihle et al., 2012; Landsiedel & Gilbert, 2015; Marsh et al., 1998; Maylor, 1990; Scarampi & Gilbert, 2020). We also replicated previous research demonstrating that higher working memory ability was associated with better memory for delayed intentions without offloading (e.g., Brewer et al., 2010). Critically, however, this relation was eliminated with reminders. Although working memory was not associated with optimal offloading decisions, those with lower ability did choose to offload more often. These findings suggest that offloading is a fruitful method to reduce memory for delayed intentions failures and that individuals with poor

2

cognitive ability may benefit most from doing so. Below we discuss the theoretical and applied ramifications for these findings.

The multi-component model of working memory suggests that in addition to the overall capacity, individual differences in working memory are driven by two components: an attention component to maintain goal-relevant information and a memory component to retrieve information from long-term memory (Unsworth et al., 2014). We have argued that the same processes that underlie working memory are needed to notice targets and remember the contents of the intention (Ball et al., 2019), as those with lower working memory typically do worse on prospective memory tasks that place high demands on attention (Arnold et al., 2015; Ball & Brewer, 2018; Ball et al., 2013; Ball et al., 2019; Brewer et al., 2010; Smith & Bayen, 2005; Unsworth, Brewer, et al., 2012) or memory (Ball et al., 2013; Ball et al., 2018). In the current set of tasks, participants not only had to notice that a target required action (i.e., attention), but also remember the location (e.g., memory) to which it should ultimately be dragged. Trying to coordinate these goals with multiple intentions while busily engaged in ongoing activities can be difficult. We replicated previous research showing that individuals with poor working memory ability performed considerably worse without offloading, presumably due to difficulties in using internal attention and memory stores to complete the tasks. Critically, the ability to externalize these demands onto the environment by immediately dragging targets near their correct location resulted in drastic improvements, with memory for delayed intentions increasing from approximately 60% without offloading to over 90% with offloading. This change is quite astounding and suggests that offloading can show drastic reductions in memory failures across a range of different tasks. Critically, the benefit of reminders was greater for individuals with poor working memory ability who have difficulty managing

prospective memory demands internally. This suggests that offloading demands onto the environment can circumvent capacity limitations.

Gilbert et al. (2020) argue that decisions to offload might occur for at least two reasons, including underconfidence in one's own memory ability and desire to avoid relying on effortful internal processing. Replicating previous research, we found that participants offloaded at lower point values than was optimal based on their objective performance (i.e., reminder bias) and were generally underconfident in their performance predictions (i.e., metacognitive bias; Gilbert et al., 2020). Importantly, those who were more underconfident in their ability also showed a greater reminder bias at the individual task level, consistent with the metacognitive bias view. This result falls in line with previous findings that JOLs and strategy use correlate with prospective memory task performance (Kuhlmann, 2019; Rummel et al., 2019; Schnitzspahn et al., 2011; Susser & Mulligan, 2019), and suggests that people are aware of the benefit of offloading on performance and err on the side of caution, thus employing it more frequently than they actually need. One important caveat to this interpretation is that this relation appeared to be task-specific, meaning that metacognitive bias on one task (e.g., 123 task) was not associated with a reminder bias on the other tasks (e.g., ABC and 321 tasks). This highlights the utility of using multiple assessments of the same cognitive ability and suggests that metacognitive insights into optimal decision making may not generalize to different tasks within the same paradigm. In other words, underconfidence may be task-specific rather than a generalizable trait variable, at least in the tasks used in the current study. Importantly, while optimal offloading decisions may be driven in part by metacognitive monitoring, this view does not seem to adequately capture the role of working memory in delayed intention task performance.

When examining individual differences in optimal offloading decisions, it was found that working memory was not related to reminder bias *or* metacognitive bias. This was somewhat surprising given that prior research has found that low ability participants engage in less effective strategy selection (Gonthier & Thomassin, 2015) and have poorer metacognitive insights into their performance (Touron et al., 2010). This perhaps makes some sense in light of previous research showing that low working memory participants (Ball et al., 2013) and older adults (Scarampi & Gilbert, 2020) can sometimes be *overconfident* in their prospective memory abilities. Apparently, low ability participants do not account for, or are not aware of, their poorer working memory ability when deciding whether offloading would be the most efficacious way to maximize points in these tasks. It is also possible that the point system was not sufficiently motivating to encourage participants to behave optimally, although prior research certainly indicates that participants prioritize learning for information that is arbitrarily assigned higher point values (Castel et al., 2011; Stefanidi et al., 2018). In contrast, working memory was associated with the overall *proportion* of trials that were offloaded, with low ability participants choosing to offload on a greater proportion trials than high ability participants. This suggests that these participants did not strategically use point values to maximize performance, but rather they chose to offload because it would result in the greatest number of items remembered with the least amount of effort expended. This suggests that the better memory for delayed intentions by individuals with higher working memory, at least in part, reflects that they were more willing or able to complete the tasks by relying on internal memory representations when required. These findings are most consistent with effort avoidance view, such that individuals with poor working memory opt to externalize cognition to minimize effort to complete the task. Future work using disincentivizing offloading by using performance-contingent rewards (e.g., Sachdeva & Gilbert, 2020) or increasing the

difficulty of setting reminders may better disentangle the mechanistic account of how working memory guides optimal offloading decisions.

The finding that offloading eliminated differences in delayed intention task performance due to working memory ability and that working memory was associated with offloading proportion is somewhat at odds with a recent study by Morrison and Richmond (2020). As described previously, they found that in a short-term memory task that individuals with lower working memory did not benefit more from offloading and did not choose to offload more often (but see Risko & Dunn, 2015). Other than the obvious difference in domains (i.e., prospective vs. retrospective memory), the most notable difference across studies is that Morrison and Richmond allowed participants to freely choose when to engage in offloading and did not include forced external trials. If participants in their tasks were unaware of the effectiveness of offloading or were uncalibrated to their own memory ability, allowing the choice to freely offload may have limited the overall offloading utility for low ability participants. In contrast, in the current study participants were required to move each prospective target within a trial any time reminders were used. Moreover, our primary measure of interest was performance on forced internal versus external trials, rather than on choice trials. Requiring participants to offload may have been particularly beneficial for low ability participants who may have otherwise opted to offload some targets, but not others. Future research in both prospective and retrospective domains could vary how offloading occurs (e.g., within versus across trials, forced versus choice) to better understand how offloading decisions are made for individuals with differing cognitive ability.

Related to the idea that forced external trials may be particularly helpful, performance in the current study was near ceiling (95%) when offloading was allowed, regardless of ability ((see also Gilbert et al., 2020; Sachdeva & Gilbert, 2020). The lack of variability in the current study

may naturally attenuate the correlation with working memory, and so the results should be interpreted with caution. Similar high rates of remembering are also often seen in more standard "focal" laboratory prospective memory tasks (Uttl, 2008). While this may not be ideal from an individual differences perspective, it makes determining actual and optimal indifference points straight forward. Moreover, working memory differences resulted in theoretically important changes in offloading decisions, specifically in the proportion of trials that were offloaded (which was not on ceiling). We would also argue that the correlation was reasonably strong without reminders and far from ceiling, making it even *more* impressive that such a simple intervention could bring performance to unity between low and high ability participants. These improvements are particularly important given that even a single failure (e.g., taking medication) can have profound effects on health or quality of living. Future studies using more difficult tasks (or less effective reminders) may help mitigate concerns associated with high performance.

Finally, it should be noted that the delay used in the current tasks was much shorter than more standard laboratory prospective memory tasks. Standard prospective memory tasks typically include a delay or storage interval to ensure the intention is moved into and later retrieved from long-term memory rather than simply maintained in working memory (e.g., Brandimonte et al., 2001; Graf & Uttl, 2001; Kliegel et al., 2011). Although we believe that similar processes are operating in the delayed intention task and standard laboratory prospective memory tasks, we do not wish to say that they are isomorphic. Clearly, working memory is predictive of internal memory in both the delayed intention and standard tasks (Ball et al., 2019). We have also found that offloading during standard laboratory tasks improves prospective remembering, especially when demands on internal processing are sufficiently high (Peper et al., 2021). However, it remains an open question whether the results from the current study of whether the observed

patterns in the present study would replicate in a standard prospective memory paradigm with longer delays.

Conclusions

The present study found that reminders improve memory for delayed intentions, especially for people with low working memory ability. The role of metacognition in intention offloading between high and low ability individuals was less clear and deserves future research. Given the importance of coordinating and managing multiple intentions in academic (e.g., assignments, examinations, and their due dates), workplace (e.g., meetings, tasks, and deadlines), and everyday settings (e.g., medications, appointments, and social engagements), the present findings provide clear evidence for the widespread benefit of intention offloading for allowing low ability individuals to perform at a level approaching equal to their higher ability peers. Advances in technology (e.g., smartphone and computerized reminders) make offloading progressively easier to implement and reduce the cost associated with the use of reminders. Schools and workplaces would do well to invest in tools and use education for effective offloading that teaches application and metacognitive awareness to maintain a balance of optimal reminder usage.

References

Arnold, N. R., Bayen, U. J., & Smith, R. E. (2015). Hierarchical Multinomial Modeling Approaches: An Application to Prospective Memory and Working Memory. *Exp Psychol, 62*(3), 143-152. https://doi.org/10.1027/1618-3169/a000287

Ball, B. H., & Brewer, G. A. (2018, May). Proactive control processes in event-based prospective memory: Evidence from intraindividual variability and ex-Gaussian analyses. *J Exp Psychol Learn Mem Cogn, 44*(5), 793-811. https://doi.org/10.1037/xlm0000489

Ball, B. H., Knight, J. B., Dewitt, M. R., & Brewer, G. A. (2013). Individual differences in the delayed execution of prospective memories. *Q J Exp Psychol (Hove), 66*(12), 2411-2425. https://doi.org/10.1080/17470218.2013.785576

Ball, B. H., Pitaes, M., & Brewer, G. A. (2018, Sep). Individual differences in episodic memory abilities predict successful prospective memory output monitoring. *Memory, 26*(8), 1159-1168. https://doi.org/10.1080/09658211.2018.1436180

Ball, H., Vogel, A., & Brewer, G. A. (2019). Individual differences in prospective memory. *Prospective Memory*, 116-134. https://doi.org/10.4324/9781351000154-8

Boldt, A., & Gilbert, S. J. (2019, Dec 2). Confidence guides spontaneous cognitive offloading. *Cogn Res Princ Implic, 4*(1), 45. https://doi.org/10.1186/s41235-019-0195-y

Brandimonte, M. A., Ferrante, D., & Feresin, C. (2001). Dissociating prospective memory from vigilance processes. 97-113.

Brewer, G. A., Knight, J. B., Marsh, R. L., & Unsworth, N. (2010). Individual differences in event-based prospective memory: Evidence for multiple processes supporting cue detection. *Memory & cognition, 38*(3), 304-311.

Castel, A. D., Humphreys, K. L., Lee, S. S., Galvan, A., Balota, D. A., & McCabe, D. P. (2011, Nov). The development of memory efficiency and value-directed remembering across the life span: a cross-sectional study of memory and selectivity. *Dev Psychol, 47*(6), 1553-1564. https://doi.org/10.1037/a0025623

Craik, F., Klix, F., & Hagendorf, H. (1986). *A functional account of age differences in memory*. In F. Klix & H. Hagendorf (Eds.), Human memory and cognitive capabilities, mechanisms, and performances. https://books.google.com/books?hl=en&lr=&id=ajglDwAAQBAJ&oi=fnd&pg=PA147&dq=A+functional+account+of+age+differences+in+memory&ots=R5CVIWJMwa&sig=vvfka4FX6LBTHmAG_3Y26mDWZR4

Craik, F., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(3), 474-479. https://doi.org/10.1037/0278-7393.13.3.474

Dunn, T. L., & Risko, E. F. (2016, Jul). Toward a Metacognitive Account of Cognitive Offloading. *Cogn Sci, 40*(5), 1080-1127. https://doi.org/10.1111/cogs.12273

Gilbert, S. J. (2015a, 2015-05-01). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology, 68*(5), 971-992. https://doi.org/10.1080/17470218.2014.972963

Gilbert, S. J. (2015b, 2015-05-01). 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. (2015c, May). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Conscious Cogn, 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, Mar). Optimal use of reminders: Metacognition, effort, and cognitive offloading. *J Exp Psychol Gen, 149*(3), 501-517. https://doi.org/10.1037/xge0000652

Gonthier, C., & Thomassin, N. (2015, Oct). Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices. *J Exp Psychol Gen, 144*(5), 916-924. https://doi.org/10.1037/xge0000101

Graf, P., & Uttl, B. (2001, Dec). Prospective memory: a new focus for research. *Conscious Cogn, 10*(4), 437-450. https://doi.org/10.1006/ccog.2001.0504

Haines, S. J., Randall, S. E., Terrett, G., Busija, L., Tatangelo, G., McLennan, S. N., Rose, N. S., Kliegel, M., Henry, J. D., & Rendell, P. G. (2020, Jun 1). Differences in time-based task characteristics help to explain the age-prospective memory paradox. *Cognition, 202*, 104305. https://doi.org/10.1016/j.cognition.2020.104305

Ihle, A., Schnitzspahn, K., Rendell, P. G., Luong, C., & Kliegel, M. (2012). Age benefits in everyday prospective memory: the influence of personal task importance, use of reminders and everyday stress. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn, 19*(1-2), 84-101. https://doi.org/10.1080/13825585.2011.629288

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001, Jun). A controlled-attention view of working-memory capacity. *J Exp Psychol Gen, 130*(2), 169-183. https://doi.org/10.1037//0096-3445.130.2.169

Kane, M. J., & Engle, R. W. (2003, Mar). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *J Exp Psychol Gen, 132*(1), 47-70. https://doi.org/10.1037/0096-3445.132.1.47

3

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004, Jun). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *J Exp Psychol Gen, 133*(2), 189-217. https://doi.org/10.1037/0096-3445.133.2.189

Kliegel, M., Altgassen, M., Hering, A., & Rose, N. S. (2011, Jul). A process-model based approach to prospective memory impairment in Parkinson's disease. *Neuropsychologia, 49*(8), 2166-2177. https://doi.org/10.1016/j.neuropsychologia.2011.01.024

Kline, R. (2011). Principles and practice of structural equation modeling (Vol. 3rd). *New York: Guilford publications*.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Knight, J. B., Marsh, R. L., Brewer, G. A., & Clementz, B. A. (2012). Preparatory distributed cortical synchronization determines execution of some, but not all, future intentions. *Psychophysiology, 49*(9), 1155-1167. https://doi.org/10.1111/j.1469-8986.2012.01400.x

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010, Nov). Decision making and the avoidance of cognitive demand. *J Exp Psychol Gen, 139*(4), 665-682. https://doi.org/10.1037/a0020198

Kuhlmann, B. G. (2019). Metacognition of prospective memory: Will I remember to remember? In J. Rummel & M. A. McDaniel (Eds.), *Prospective Memory* (pp. 60-77). Routledge.

Landsiedel, J., & Gilbert, S. J. (2015, Jan 1). Creating external reminders for delayed intentions: dissociable influence on "task-positive" and "task-negative" brain networks. *Neuroimage, 104*, 231-240. https://doi.org/10.1016/j.neuroimage.2014.10.021

Loft, S., Smith, R. E., & Bhaskara, A. (2011). Prospective memory in an air traffic control simulation: External aids that signal when to act. *Journal of Experimental Psychology: Applied, 17*(1), 60-70. https://doi.org/10.1037/a0022845

Marsh, R. L., & Hicks, J. L. (1998, Mar). Event-based prospective memory and executive control of working memory. *J Exp Psychol Learn Mem Cogn, 24*(2), 336-349. https://doi.org/10.1037//0278-7393.24.2.336

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005, Jan). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. *J Exp Psychol Learn Mem Cogn, 31*(1), 68-75. https://doi.org/10.1037/0278-7393.31.1.68

Marsh, R. L., Hicks, J. L., & Landau, J. D. (1998, Jul). An investigation of everyday prospective memory. *Mem Cognit, 26*(4), 633-643. https://doi.org/10.3758/bf03211383

Maylor, E. A. (1990, Aug). Age and prospective memory. *Q J Exp Psychol A, 42*(3), 471-493. https://doi.org/10.1080/14640749008401233

McDaniel, M. A., Lamontagne, P., Beck, S. M., Scullin, M. K., & Braver, T. S. (2013). Dissociable Neural Routes to Successful Prospective Memory. *Psychological Science, 24*(9), 1791-1800. https://doi.org/10.1177/0956797613481233

Meeks, J. T., Hicks, J. L., & Marsh, R. L. (2007, 2007/12/01/). Metacognitive awareness of event-based prospective memory. *Consciousness and Cognition, 16*(4), 997-1004. https://doi.org/https://doi.org/10.1016/j.concog.2006.09.005

Morrison, A. B., & Richmond, L. L. (2020, Jan 3). Offloading items from memory: individual differences in cognitive offloading in a short-term memory task. *Cogn Res Princ Implic, 5*(1), 1. https://doi.org/10.1186/s41235-019-0201-4

Nelson, T., Narens, L., & Bower, G. (1990). The psychology of learning and motivation. *Metamemory: A theoretical framework and new findings*.

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*(4), 1343-1355.

Penningroth, S. L., & Scott, W. D. (2013, 2013/09/01). Task Importance Effects on Prospective Memory Strategy Use [https://doi.org/10.1002/acp.2945]. *Applied Cognitive Psychology, 27*(5), 655-662. https://doi.org/https://doi.org/10.1002/acp.2945

Peper, P., Alakbarova, D., & Ball, B. H. (2021). Differential Benefits of Prospective Memory Reminders Depending on Cognitive Load.

Reynolds, J. R., West, R., & Braver, T. (2009, May). Distinct neural circuits support transient and sustained processes in prospective memory and working memory. *Cereb Cortex, 19*(5), 1208-1221. https://doi.org/10.1093/cercor/bhn164

Risko, E. F., & Dunn, T. L. (2015, Nov). Storing information in-the-world: Metacognition and cognitive offloading in a short-term memory task. *Conscious Cogn, 36*, 61-74. https://doi.org/10.1016/j.concog.2015.05.014

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

Rummel, J., Danner, D., & Kuhlmann, B. G. (2019). The short version of the Metacognitive Prospective Memory Inventory (MPMI-s): factor structure, reliability, validity, and reference data. *Measurement Instruments for the Social Sciences, 1*(1). https://doi.org/10.1186/s42409-019-0008-6

Sachdeva, C., & Gilbert, S. J. (2020). Excessive use of reminders: Metacognition and effort-minimisation in cognitive offloading. *Consciousness and Cognition, 85*, 103024.

Scarampi, C., & Gilbert, S. J. (2020, May). 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, Mar). Age differences in strategic reminder setting and the compensatory role of metacognition. *Psychol Aging, 36*(2), 172-185. https://doi.org/10.1037/pag0000590

Schnitzspahn, K. M., Zeintl, M., Jager, T., & Kliegel, M. (2011, Mar). Metacognition in prospective memory: are performance predictions accurate? *Can J Exp Psychol, 65*(1), 19-26. https://doi.org/10.1037/a0022842

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609-612.

Shelton, J. T., Scullin, M. K., & Hacker, J. Y. (2019). The multiprocess framework: Historical context and the "dynamic" extension. In J. Rummel & M. A. McDaniel (Eds.), *Prospective Memory* (pp. 10-26). Routledge.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017, Jul 25). Toward a Rational and Mechanistic Account of Mental Effort. *Annu Rev Neurosci, 40*, 99-124. https://doi.org/10.1146/annurev-neuro-072116-031526

Smith, R. E., & Bayen, U. J. (2005). The effects of working memory resource availability on prospective memory: a formal modeling approach. *Exp Psychol, 52*(4), 243-256. https://doi.org/10.1027/1618-3169.52.4.243

Stefanidi, A., Ellis, D. M., & Brewer, G. A. (2018). Free recall dynamics in value-directed remembering. *Journal of Memory and Language, 100*, 18-31. https://doi.org/10.1016/j.jml.2017.11.004

Susser, J. A., & Mulligan, N. W. (2019, 2019/02/01/). Exploring the intrinsic-extrinsic distinction in prospective metamemory. *Journal of Memory and Language, 104*, 43-55. https://doi.org/https://doi.org/10.1016/j.jml.2018.09.003

Touron, D. R. (2015, Jun). Memory avoidance by older adults: When `old dogs' won't perform their `new tricks'. *Curr Dir Psychol Sci, 24*(3), 170-176. https://doi.org/10.1177/0963721414563730

Touron, D. R., Oransky, N., Meier, M. E., & Hines, J. C. (2010). Metacognitive monitoring and strategic behaviour in working memory performance. *Quarterly Journal of Experimental Psychology, 63*(8), 1533-1551.

3

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2012). Variation in cognitive failures: An individual differences investigation of everyday attention and memory failures. *Journal of Memory and Language, 67*(1), 1-16. https://doi.org/10.1016/j.jml.2011.12.005

Unsworth, N., & Engle, R. W. (2007, Jan). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol Rev, 114*(1), 104-132. https://doi.org/10.1037/0033-295X.114.1.104

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014, Jun). Working memory and fluid intelligence: capacity, attention control, and secondary memory retrieval. *Cogn Psychol, 71*, 1-26. https://doi.org/10.1016/j.cogpsych.2014.01.003

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004, Nov). Working memory capacity and the antisaccade task: individual differences in voluntary saccade control. *J Exp Psychol Learn Mem Cogn, 30*(6), 1302-1321. https://doi.org/10.1037/0278-7393.30.6.1302

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012, 2012-12-01). Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *Quarterly Journal of Experimental Psychology, 65*(12), 2397-2410. https://doi.org/10.1080/17470218.2012.690438

Uttl, B. (2008, Feb 20). Transparent meta-analysis of prospective memory and aging. *PLoS One, 3*(2), e1568. https://doi.org/10.1371/journal.pone.0001568

Correlations of intention memory and working memory measures

	PML_Internal_123	PML_Internal_ABC	PML_Internal_321	PML_External_123	PML_External_ABC	PML_External_321	WMC_Ospan	WMC_Rspan	WMC_Sspan
PML_Internal_123		0.477***	0.402***	0.313***	0.227***	0.1	0.187**	0.15*	0.317***
PML_Internal_ABC	0.477***		0.474***	0.246***	0.216***	0.147*	0.207***	0.181**	0.223***
PML_Internal_321	0.402***	0.474***		0.182**	0.115	0.177**	0.1	0.122*	0.262***
PML_External_123	0.313***	0.246***	0.182**		0.252***	0.235***	-0.055	-0.072	0.087
PML_External_ABC	0.227***	0.216***	0.115	0.252***		0.216***	0.117	0.097	0.14*
PML_External_321	0.1	0.147*	0.177**	0.235***	0.216***		-0.001	0.047	0.143*
WMC_Ospan	0.187**	0.207***	0.1	-0.055	0.117	-0.001		0.61***	0.312***
WMC_Rspan	0.15*	0.181**	0.122*	-0.072	0.097	0.047	0.61***		0.37***
WMC_Sspan	0.317***	0.223***	0.262***	0.087	0.14*	0.143*	0.312***	0.37***	

Variables (columns/rows): PML_Internal_123, PML_Internal_ABC, PML_Internal_321, PML_External_123, PML_External_ABC, PML_External_321, WMC_Ospan, WMC_Rspan, WMC_Sspan

	PML_Internal_123	PML_Internal_ABC	PML_Internal_321	PML_External_123	PML_External_ABC	PML_External_321	WMC_Ospan	WMC_Rspan
PML_Internal_ABC	$r = 0.477$							
PML_Internal_321	$r = 0.402$	$r = 0.474$						
PML_External_123	$r = 0.313$	$r = 0.246$	$r = 0.182$					
PML_External_ABC	$r = 0.227$	$r = 0.216$	$r = 0.115$	$r = 0.252$				
PML_External_321	$r = 0.100$	$r = 0.147$	$r = 0.177$	$r = 0.235$	$r = 0.216$			
WMC_Ospan	$r = 0.187$	$r = 0.207$	$r = 0.100$	$r = -0.055$	$r = 0.117$	$r = -0.001$		
WMC_Rspan	$r = 0.150$	$r = 0.181$	$r = 0.122$	$r = -0.072$	$r = 0.097$	$r = 0.047$	$r = 0.610$	
WMC_Sspan	$r = 0.317$	$r = 0.223$	$r = 0.262$	$r = 0.087$	$r = 0.140$	$r = 0.143$	$r = 0.312$	$r = 0.370$
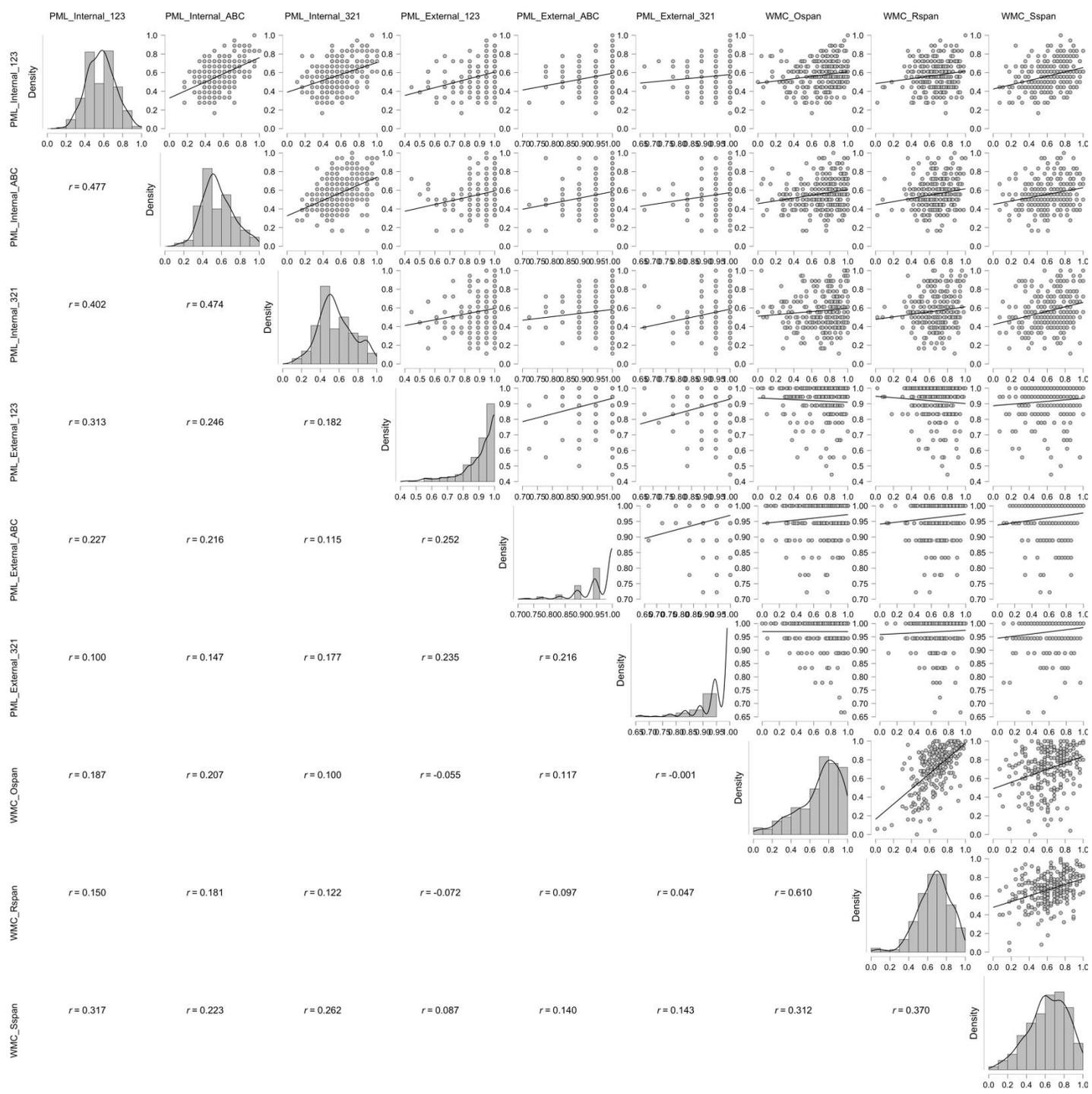
Correlations of offloading proportion and working memory measures
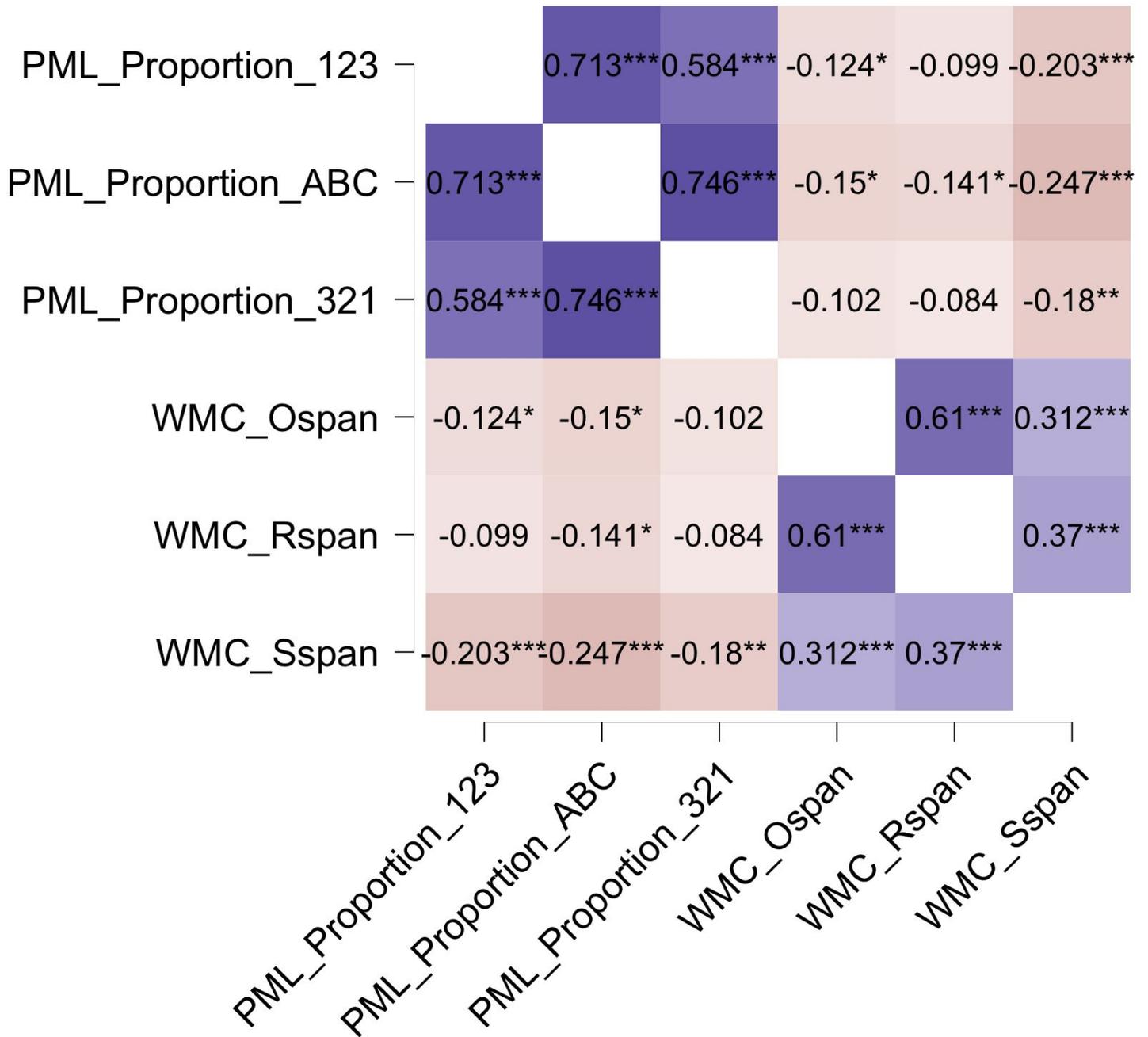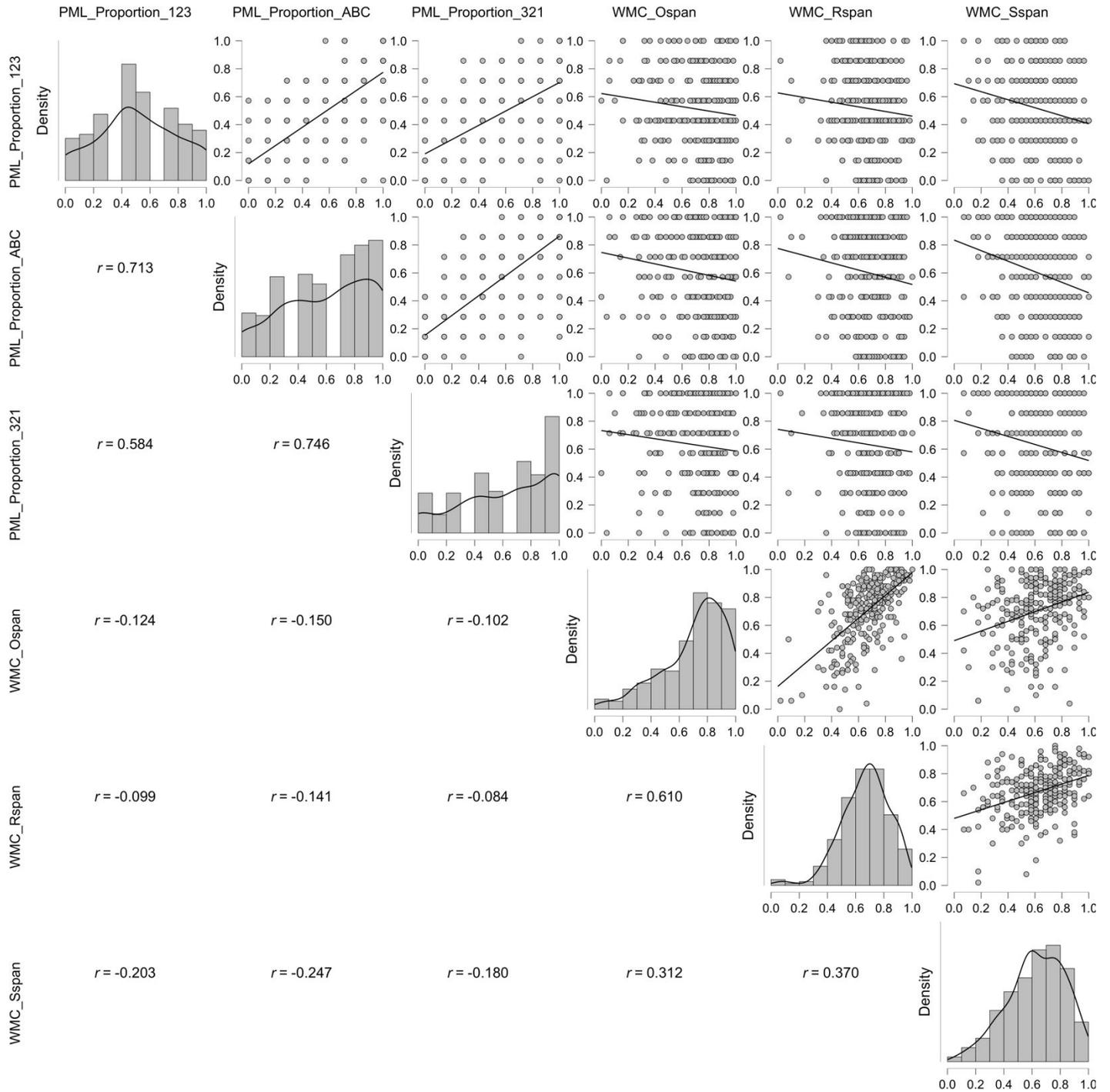
Correlations of biases (reminder and metacognitive) and working memory measures

	PML_Reminder_Bias_123	PML_Reminder_Bias_ABC	PML_Reminder_Bias_321	PML_Metacognitive_Bias_123	PML_Metacognitive_Bias_ABC	PML_Metacognitive_Bias_321	WMC_Ospan	WMC_Rspan	WMC_Sspan
PML_Reminder_Bias_123		0.437***	0.414***	-0.219***	0.109	0.133*	0.007	0.011	-0.002
PML_Reminder_Bias_ABC	0.437***		0.554***	-0.035	-0.173**	-0.006	-0.01	-0.014	-0.053
PML_Reminder_Bias_321	0.414***	0.554***		-0.018	0.029	-0.176**	-0.031	-0.025	-0.027
PML_Metacognitive_Bias_123	-0.219***	-0.035	-0.018		0.348***	0.32***	0.033	-0.005	-0.101
PML_Metacognitive_Bias_ABC	0.109	-0.173**	0.029	0.348***		0.602***	-0.037	-0.016	-0.066
PML_Metacognitive_Bias_321	0.133*	-0.006	-0.176**	0.32***	0.602***		0.003	0.008	-0.132*
WMC_Ospan	0.007	-0.01	-0.031	0.033	-0.037	0.003		0.61***	0.312***
WMC_Rspan	0.011	-0.014	-0.025	-0.005	-0.016	0.008	0.61***		0.37***
WMC_Sspan	-0.002	-0.053	-0.027	-0.101	-0.066	-0.132*	0.312***	0.37***	

PML_Reminder_Bias_123　PML_Reminder_Bias_ABC　PML_Reminder_Bias_321　PML_Metacognitive_Bias_12　ML_Metacognitive_Bias_AB　PML_Metacognitive_Bias_32　WMC_Ospan　WMC_Rspan　WMC_Sspan

	PML_Reminder_Bias_123	PML_Reminder_Bias_ABC	PML_Reminder_Bias_321	PML_Metacognitive_Bias_12	ML_Metacognitive_Bias_AB	PML_Metacognitive_Bias_32	WMC_Ospan	WMC_Rspan
PML_Reminder_Bias_ABC	$r = 0.437$							
PML_Reminder_Bias_321	$r = 0.414$	$r = 0.554$						
PML_Metacognitive_Bias_12	$r = -0.219$	$r = -0.035$	$r = -0.018$					
PML_Metacognitive_Bias_32	$r = 0.109$	$r = -0.173$	$r = 0.029$	$r = 0.348$				
PML_Metacognitive_Bias_32	$r = 0.133$	$r = -0.006$	$r = -0.176$	$r = 0.320$	$r = 0.602$			
WMC_Ospan	$r = 0.007$	$r = -0.010$	$r = -0.031$	$r = 0.033$	$r = -0.037$	$r = 0.003$		
WMC_Rspan	$r = 0.011$	$r = -0.014$	$r = -0.025$	$r = -0.005$	$r = -0.016$	$r = 0.008$	$r = 0.610$	
WMC_Sspan	$r = -0.002$	$r = -0.053$	$r = -0.027$	$r = -0.101$	$r = -0.066$	$r = -0.132$	$r = 0.312$	$r = 0.370$