In Press, Consciousness and Cognition

Registered Report

Excessive use of reminders: Metacognition and effort-minimisation in cognitive offloading

Chhavi Sachdeva & Sam J. Gilbert

Institute of Cognitive Neuroscience, University College London, UK

Keywords: prospective memory; intentions; metacognition; cognitive offloading; effort; reward

Address correspondence to:

Chhavi Sachdeva
Institute of Cognitive Neuroscience
17 Queen Square
London WC1N 3AZ
UK
Email: chhavi.sachdeva.14@ucl.ac.uk
Tel: +44 (0)20 7679 1199
Fax: +44 (0)20 7813 2835

Acknowledgements: Supported by the Economic and Social Research Council (ESRC) grant ES/N018621/1

People often use external reminders to help remember delayed intentions. This is a form of "cognitive offloading". Individuals sometimes offload more often than would be optimal (Gilbert et al., 2020). This bias has been linked to participants' erroneous metacognitive underconfidence in their memory abilities. However, underconfidence is unlikely to fully explain the bias. An additional, previously-untested factor that may contribute to the offloading bias is a preference to avoid cognitive effort associated with remembering internally. The present experiment examined evidence for this hypothesis. One group of participants received payment contingent on their performance of the task (hypothesised to increase cognitive effort, and therefore reduce the bias towards offloading); another group received a flat payment for taking part, as in the earlier experiment. The offloading bias was significantly reduced (but not eliminated) in the rewarded group, suggesting that a preference to avoid cognitive effort influences cognitive offloading.

Introduction

In everyday life, intentions often need to be remembered and performed after a delay. For example, your doctor might prescribe you with antibiotics that need to be taken at regular intervals. You might choose to remember this by maintaining the intention internally. Alternatively, you could create an external reminder by placing the antibiotics next to your bed, setting an alert on your smartphone, or asking "Alexa" or "Siri" to remind you. These are examples of "cognitive offloading" (Risko & Gilbert, 2016) – the use of physical action to reduce the cognitive demands of a task. The use of cognitive offloading to supplement memory is increasingly pervasive due to the rapid development of technology. It is therefore important to understand when and how people rely on offloading strategies so that individuals can be guided towards more effective use of cognitive tools to maximise their memory performance.

Recent research has started to investigate how people decide between relying on their internal resources (i.e., remembering the intention using their own memory) and external resources (e.g., setting reminders on external artefacts) in order to remember delayed intentions (e.g., Gilbert, 2015a, 2015b; Gilbert et al., 2020). One possible reason for using an external reminder is a belief that one might otherwise forget. This implies that the choice to offload is driven by an attempt to maximise accuracy. An alternative reason is to reduce the effort associated with remembering an intention internally. This implies that the choice to offload is driven by an attempt to minimise cognitive effort. These alternatives are conceptually distinct, but not mutually exclusive. The main aim of this study is to investigate the role of these two factors in cognitive offloading.

Previous studies in offloading: The role of metacognition

A paradigm studying cognitive offloading was developed by Gilbert (2015a). In this task, participants were presented with a set of numbered circles which they had to drag sequentially to the bottom of the screen with their computer mouse. Before beginning each trial, they were instructed to drag one or more of the circles to an alternative location (e.g., drag “4” to the left). Participants could either remember this intention internally, or they could “offload” the intention by placing the target circle next to its corresponding side of the screen as soon as the trial began. An everyday analogy would be placing an object by the front door so that you remember to bring it with you when leaving home the next morning. In this study, participants were given a free choice between setting external reminders or remembering the intentions internally. They were more likely to offload intentions when they had more items to remember, or when they encountered interruptions during the ongoing task. Both of these factors reduced accuracy when offloading was not permitted, suggesting that the decision whether or not to set reminders was based, at least in part, on participants’ evaluation of the difficulty of the task.

Gilbert (2015b) followed this up by asking participants to report their confidence in their ability to perform the task, as well as measuring their objective memory ability, before introducing the cognitive offloading strategy. Participants who had lower confidence in their memory abilities were more likely to set reminders, regardless of their objective memory ability. This was replicated by Boldt & Gilbert (2019) both when the reminder-setting strategy was explicitly explained to participants and when they had to spontaneously generate it.

These findings highlight the importance of metacognitive processes in reminder-setting behaviour. Metacognition refers to a range of processes involved in reflecting on or evaluating one’s mental states, including the ability to estimate confidence in one’s performance of a task (Fleming et al., 2012; Yeung & Summerfield, 2012). Results from a range of paradigms indicate that one of the factors contributing to cognitive offloading behaviour is a belief that performance will be poor otherwise, regardless of objective cognitive ability (see Risko & Dunn, 2015; Risko & Gilbert, 2016).

Biases and optimality in cognitive offloading

Setting a reminder incurs a cost (interrupting the ongoing task, and the time it takes to set a reminder) and a benefit (increased likelihood of remembering). Gilbert et al. (2020) developed an experimental paradigm to investigate whether participants optimally balance such costs and benefits. Participants performed a difficult task in which accuracy was low (approximately 50%) when using internal memory, but close to 100% when using external reminders. They were then given a series of choices between earning a maximum reward using internal memory alone (10 points per remembered item), or a smaller reward (1-9 points) with reminders allowed. This allowed the optimality of choice behaviour to be examined. For example, if a participant can achieve 55% accuracy using internal memory and 100% accuracy using reminders, it would be optimal to choose internal memory when offered 5 points per item with reminders, and external reminders when offered 6 points per item. Gilbert et al. (2020) found that participants were systematically biased, tending to choose reminders even when they would have earned more points with internal memory. This “reminder bias” was observed both with and without a financial incentive to choose optimally, and individual differences in bias were stable over time. However, the effect of financial incentive was not directly investigated by manipulating this factor within a single experiment. Gilbert et al. (2020) also found that bias was correlated with participants’ “metacognitive bias”, i.e. the discrepancy between confidence in their internal memory ability and objective performance. That is, participants who believed their internal memory to be poor were more likely to choose reminders, regardless of their objective memory performance. Furthermore, Gilbert et al. (2020 experiment 3) investigated the effect of two interventions designed to influence metacognitive judgements ("metacognitive interventions"). Both interventions shifted the reminder bias, and these shifts were mediated by shifts in confidence.

The two metacognitive interventions investigated by Gilbert et al. (2020) were feedback valence (whether participants received positive or negative feedback about their performance) and initial practice difficulty (whether participants received an easier version of the task at the beginning or more difficult one). These interventions were crossed in a 2x2 design leading to four groups of participants. Negative feedback led to reduced confidence and increased use of reminders, as did difficult practice. In three of the four groups, participants were underconfident, i.e. their predicted accuracy was below objective accuracy. However, participants in the easy-practice, positive-feedback group were overconfident. Despite this, all four groups were significantly biased towards excessive use of reminders. If metacognitive judgements were the only factor related to reminder bias, one would expect underconfident individuals to use too many reminders and overconfident individuals to use too few. However, results of Gilbert et al. (2020) show that reminder bias can be observed both in the context of under- and over-confidence. How might this be explained?

One possibility is that the metacognitive measure was inaccurate. Gilbert et al. (2020), used only one metacognitive judgement scale, which was presented to participants after the completion of the practice trials. It is possible that as the experimental trials progressed, participants became increasingly underconfident in their performance as the distance from the interventions used was large enough to stop influencing behaviour. Consistent with this possibility, West and Mulligan (2019) demonstrated that prospective metamemory (i.e., confidence in one's prospective memory abilities), like retrospective metamemory (i.e., confidence in one's memory about past events or experiences) displays underconfidence with practice, meaning that, the more practice participants get, the more underconfident they become in their cognitive abilities. This is known as the underconfidence with practice (UWP) effect (Koriat et al., 2002). The UWP effect might explain the pattern of biases observed in the easy, positive group, where participants might have been overconfident at the beginning of the experiment, but slowly become increasingly underconfident as they gained more practice. To evaluate this possibility, the current experiment will include a second metacognitive judgement scale at the end of the experiment to investigate whether there is a change in participants' confidence from the beginning to the end of the experiment.

Cognitive effort and reward

A second possible explanation for the findings of Gilbert et al. (2020 experiment 3) is that, in addition to the metacognitive bias, there might be one or more additional factors that contribute to the reminder bias. One such potential factor is a preference to avoid cognitive effort. The concept of cognitive effort has proved difficult to define mechanistically (Shenhav et al., 2017). Instead, effortful tasks are typically defined in terms of their phenomenology of being subjectively difficult or demanding, or in terms relatively poor performance (i.e. low accuracy and/or high response time; see Gilbert, Bird, Frith, & Burgess, 2012 for discussion on the concept of task difficulty). Research has suggested that effort is aversive (Dreisbach & Fischer, 2015; Kurzban, 2016; Saunders et al., 2017; Shenhav et al., 2017) and individuals tend to avoid effortful tasks (Frederick, 2005; Kool et al., 2010). Often called the "law of less work", when given a choice between similarly rewarding options, organisms typically learn to avoid the option that requires more work or effort (Hull, 1943).

The argument that individuals tend to avoid cognitive effort is consistent with the view that they have an intrinsic drive to avoid internal memory resources, and instead use external perceptual information instead (Ballard, Hayhoe, Pook, & Rao, 1997). However, there is evidence that cognitive effort is not costly in all circumstances and can even be rewarding (Eisenberg, 1992; Inzlicht, Shenhav, & Olivola, 2018). Consistent with this, individuals can sometimes show a bias towards the use of internal rather than external resources (Walsh and Anderson, 2009).

One theoretical account of subjective effort suggests that effortful activities involve cognitive processes (such as those associated with working memory) which are both limited in capacity and potentially applicable to a wide range of activities (Kurzban et al., 2013). Therefore, individuals will tend to avoid the expenditure of cognitive effort, so that these processes can be redirected towards other activities. This could explain why remembering an intention internally feels more effortful than setting an external reminder: to the extent that internal memory capacity is occupied by currently-active intentions, this precludes the use of that capacity for other purposes. By contrast, once an external reminder has been set, this exerts no detrimental effect on our ability to pursue other activities.

To further examine the role of effort in cognitive offloading, the current experiment tested whether a bias towards reminders can be explained by effort-avoidance, by manipulating financial incentive through performance-based rewards. We hypothesised that a financial incentive would provide participants with more motivation to expend cognitive effort. Research in different domains has suggested that rewards give individuals incentives to work harder (see Aarts et al., 2010 for evidence in selective attention; see Padmala & Pessoa, 2011 for evidence in task switching). Since incentives motivate individuals to work harder, they are regularly administered to improve cognitive performance (Botvinick & Braver, 2015). For example, Krebs, Boehler and Woldorff (2010) found that participants responded faster and more accurately when expecting greater reward for naming the colour of a Stroop stimulus. The notion that individuals are willing to expend more effort when rewards are available is known as motivational vigor (Berridge, 2004; Niv et al., 2006). In support of this idea, research has found that the prospect of performance-contingent reward promotes cognitive stability and proactive control (Fröber & Dreisbach, 2014; Jimura et al., 2010; Locke & Braver, 2008; Padmala & Pessoa, 2011). Furthermore, in their experiment 6B Kool, McGuire, Rosen and Botvinick (2010) used a paradigm in which participants repeatedly chose between two visual stimuli. Each switch between the two stimuli was classified as either high demand or low demand. They found that participants constantly gravitated towards the low demand option, but this bias was reduced (although not eliminated) when a monetary incentive was linked with the high-effort option.

The evidence outlined suggests that financial incentives increase effort in tasks. So, if one contributor to the reminder bias is that people set reminders to avoid effort, financial incentive should reduce this bias. It should also be pointed out, however, that individuals might still have an intrinsic bias against cognitive effort which is not fully compensated by the reward offered. Indeed, Westbrook, Kester and Braver (2013) found that participants would accept a financial penalty to perform a task that is less cognitively demanding. Therefore, although we predict that a financial incentive could reduce the reminder bias, it may not eliminate it.

Current research

The current experiment was a replication of experiment 3 from Gilbert et al. (2020), adapted slightly to manipulate performance-based rewards. Of the four groups in the earlier experiment, only the easy-practice, positive-feedback group simultaneously showed overconfidence and a positive reminder bias (i.e. a reminder bias that cannot be explained in terms of metacognitive error). Therefore, only this condition was replicated here. One group of participants received a base payment without any bonus financial incentives (as in Gilbert et al., 2020 experiment 3) while the other group received performance-based rewards in addition to the base payment. Furthermore, a second metacognitive judgement scale (measuring participants' confidence), was included at the end of the experiment to investigate whether there was a change in participants' confidence from the beginning to the end of the experiment.

Methods

Design

This experiment followed the procedure of the "easy-practice, positive-feedback" condition of Gilbert et al. (2020 experiment 3), with some changes that will be highlighted. Participants were randomly assigned to one of two groups, reward versus no reward. Participants in the reward group received payment based on the number of points they scored in the task, while those in the no reward group received a fixed payment regardless of performance. Thus, participants in the no reward group fully replicated the procedure of Gilbert et al. (2020 experiment 3). Data from this group was compared with a group that received a performance-based payment. In addition, both groups provided an additional confidence judgement at the end of the experiment, which was not collected in the earlier study.

Power calculations

We aimed to address three key questions in this study:

a) Can we replicate our earlier findings from the “easy-practice, positive-feedback” group, that participants are both overconfident in their internal memory abilities and have a bias towards external reminders? This was initially examined in the no-reward group only, seeing as this group replicated the earlier procedure. The previous effect sizes (Cohen’s d) for the metacognitive bias and reminder bias were 0.34 and 0.44 respectively. In order to achieve 90% power to detect effects of these sizes (one-sample one-tailed t tests) we required 76 and 46 participants respectively (G*Power 3.1). Seeing as these analyses applied to just one of the two groups, this translated to a total sample of 152 and 92 respectively, assuming equal numbers of participants in each group. Therefore, a total of 152 participants is required for 90% power to detect the smaller effect size.

b) Is the reminder bias of the no-reward group reduced in the reward group? In the previous experiment, participants showed a positive reminder bias (mean = 1.2, SD = 2.7) despite being overconfident. If financial incentive removed any bias against cognitive effort, participants might be expected to have a *negative* reminder bias, i.e. a bias against using external reminders, seeing as they were overconfident in their internal memory abilities. However, the financial incentive might not entirely eliminate a bias against cognitive effort (cf. Westbrook et al., 2013). Therefore, we based our power calculation more conservatively on a scenario where the reminder bias of the reward group is reduced to zero, rather than becoming negative. Assuming that both groups would have a similar standard deviation, this implies a comparison between two groups with means 1.2 and 0, both of which have a SD of 2.7. This equates to a Cohen’s d of 0.44. To achieve 90% power to detect an effect of this size with a two-sample, one-tailed, t-test 180 participants are required (90 in each group). We used a one-tailed test seeing as we had a directional hypothesis and only tested for a difference in this direction.

c) Are participants less confident for the second metacognitive judgement (at the end of the experiment) than the first? West and Mulligan (2019) found an underconfidence with practice (UWP) effect in their prospective memory task with an effect size (η^2_p) of 0.15 (Experiment 2, comparison between Blocks 1 and 2). In order to achieve 90% power to detect an effect of this size, a sample size of 32 would be required (G*Power 3.1), under the most conservative assumption that the repeated measures are uncorrelated. We conducted a two-tailed test here, seeing as we thought it was also possible that participants could become more confident following practice, and would wish to examine any such effect statistically.

These power calculations suggest that a sample size of 180 is sufficient for adequate power to test for our smallest predicted effect. However, ensuring that the study has sufficient power to test the smallest effect alone does not guarantee sufficient power to test all hypotheses together (Francis & Thunell, 2019). Under the conservative assumption that all four analyses described above are independent (i.e. a participant producing data consistent with one hypothesis is no more likely to produce data consistent with the others), we simply multiplied the post-hoc power associated with each of our four tests to obtain the overall power to detect all four effects together. With a total sample size of 180, the power to detect all four effects was 84%. Assuming equal numbers of participants in each group, this needed to rise to 208 in order to achieve power of 90%. Therefore, we tested a total of 208 participants, with 104 in each group.

Procedure

Participants performed a task previously used by Gilbert et al. (2020), taking part via their computer's web browser (see Figure 1 for a schematic illustration). On each trial, they viewed six yellow circles randomly positioned within a square. Each circle contained a number, and participants were asked to use their computer mouse to drag the circles sequentially (in numerical 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 numerical sequence (e.g. numbers 1-6 were on the screen, after the 1 was dragged to the bottom it would be replaced with a 7). This continued until a total of 25 circles had been dragged out of the square. Occasionally, new circles initially appeared in blue, orange, or pink, rather than yellow (these were described as "special circles" in the instructions to participants). These colours corresponded with the left, top, and right side of the square respectively. Two seconds after appearing on the screen, their colour faded to yellow so that they matched the other circles. When a new circle appeared in one of these colours, this represented an instruction that it should eventually be dragged to its corresponding side of the square when it was reached in the sequence. For example, a participant drags 1 to the bottom of the screen where it disappears. An orange 7 appears in its place, fading to yellow after 2 s. Meanwhile, the participant drags circles 2-6 to the bottom of the screen, before dragging 7 to the top. Therefore, a circle temporarily appearing in a nonyellow colour instructed participants to form a delayed intention to drag that circle to a nonstandard location when it was eventually reached in the sequence. To remember this instruction, participants could either rely on an internal representation of their intention, or create an external reminder. They created external reminders by immediately dragging the target circles near their instructed location as soon as they appeared on the screen. For example, as soon as an orange

7 appeared on the screen, the participant could drag this circle to near the top of the square. Then, when they reached 7 in the sequence its location would remind the participant of their intention. In this case, there is no need to maintain an internal representation of the intended behaviour, seeing as it is directly cued by the circle's position.

One trial consisted of the numerical sequence from 1-25. Within this sequence, a total of 10 target circles appeared, randomly allocated to 10 of the numbers from 7-25. This meant that participants needed to remember multiple simultaneous intentions and it was unlikely that they would be able to remember all of them without setting external reminders. The 10 target circles were randomly allocated to the left, top, and right positions of the square. Feedback was provided as follows: When a target circle is correctly dragged to the top, left, or right side of the box, it will turn green before disappearing. When a circle is dragged to the bottom of the box, it will turn purple before disappearing regardless of whether it is a target or nontarget, which will not provide any feedback. In addition, participants received a further feedback message following completion of the 25-circle sequence, at the end of each trial (see below for details).

The full experiment can be accessed via the following weblink, which is identical to the task performed by the experimental participants:

http://www.ucl.ac.uk/sam-gilbert/CS1/DemoRM/WebTasks.html

[Figure 1 about here]

Reward manipulation

At the beginning of the experiment, participants were randomised into one of two groups: reward or no-reward. Participants in the reward group received the following set of instructions, "Your payment has not yet been determined. For this experiment, you will earn a base payment of $2.50. Additionally, you will also earn $1 for every 250 points you score. This means that you can earn up to $9.30 for this experiment". Participants in the no-reward group received the following instructions: "Your payment has now been determined. You will earn a base payment of $2.50 and an additional $5 as a bonus for taking part. This means that you will earn a total of $7.50 for completing this experiment". The reason for this phrasing was because this experiment was advertised on Amazon Mechanical Turk as an experiment paying $2.50 plus bonus. So, participants in the reward group earned a base payment of $2.50 and an additional bonus dependent on their performance. On the other hand, participants in the no-reward group earned a base payment of $2.50 and an additional predetermined bonus of $5 that was not dependent on performance. Based on the results from Gilbert et al.'s (2020 experiment 3), where participants in the easy-positive group scored an average of 1154 points, we expected this to equate to a performance-dependent bonus of $4.62 (including the $2.50 base payment, this amount would total $7.12). Therefore, even though the maximum potential reward was higher in the reward group, we expected the mean earnings to be comparable between the two groups.

Practice trials

After the above set of instructions, participants were presented with 8 practice trials. For the first two practice trials, the sequence involved a total of 7 circles with no targets. This was done so that participants could practice dragging circles to the bottom of the screen. Next, the instructions for how to respond to targets was presented and the participants performed two practice trials involving 8 circles and 1 target. They were only be able to proceed to the next phase of the experiment if they responded correctly to the target on the second of these practice trials. If they did not, another practice trial was presented.

Following this, there were five more practice trials, each with a sequence of 25 circles, of which 4 were targets. Note that this corresponded to the “easy practice” condition of Gilbert et al. (2020 experiment 3). Target circles occurred between circle numbers 7 and 25 in the sequence and were distributed in a manner that maximised (and equalised) the inter-target gap as much as possible. From this point onwards, participants received a feedback message after completing each trial, based on their target circle accuracy as follows (note that this corresponds to the “positive feedback” condition of Gilbert et al., 2020 experiment 3):

Accuracy	**Feedback**
0%	You did not get any special circles correct this time
Above 0%, below 50%	Well done – good work! You are responding well to the special circles
Above 50%, below 100%	Well done – excellent work! You responded correctly to most of the special circles
100%	Well done – perfect! You responded correctly to all of the special circles

After the 4 practice trials each containing 4 targets, participants were told, “Now the task will get more difficult. It will stay like this for the rest of the experiment. Please ignore the difficulty of the practice trials you have just done and remember that the task will be like this from now on”. They then received one final practice trial with a total of 10 targets.

Metacognitive judgement

After this, participants were asked to make a metacognitive judgement with the following instructions:

Now that you have had some practice with the experiment, we would like you to tell us how accurately you will be able to perform the task without any reminders. Please ignore the earlier practice trials and just tell us how accurately you will be able to do the task when it is the same difficulty as the trial you have just completed. The difficulty will stay the same as this for the rest of the experiment. Please use the scale below to indicate what percentage of the special circles you can correctly drag to the instructed side of the square, on average. 100% would mean that you always get every single one correct. 0% would mean that you can never get any of them correct.

They were then presented with a slider on the screen allowing them to select any percentage between 0-100% before continuing.

Intention offloading practice

After the metacognitive judgement scale, participants were presented with one practice trial that instructed them how to set reminders by dragging target circles next to their intended locations. If participants got fewer than 8 out of 10 correct on this trial, they were asked to repeat it. This was done to ensure that participants were able to achieve a high level of accuracy when using this strategy.

Scoring points

After completing the intention offloading practice, participants were told, “From now on, you will score points every time you drag one of the special circles to the correct location. You should try to score as many points as you can". Participants in the no-reward condition were then told, "Please bear in mind that points earned will not give you more money, but you should try to score as many points as you can.", while participants in the reward group were told, "You will earn a bonus depending on how many points you score. The more points you score, the more money you will earn".

When participants clicked through to the next page, they saw a red button labelled "Special circles worth 10 points. Reminders not allowed". Above the button was shown the instruction: "Sometimes when you do the task, you will have to do it without setting any reminders. When this happens, you will score 10 points for every special circle you remember. You will always be given clear instructions as to what you should do. In this case you will be told, 'This time you must do the task without setting any reminders' and you will see a red button. When this happens, the computer will not let you set any reminders. Let's practise that now".

After clicking the red button, they were presented with one practice trial in which reminders were not allowed. This was accomplished by fixing all circles on the screen apart from the next in the sequence (e.g., after you drag "2" to the bottom, only the circle labelled "3" will be moveable, ensuing that participants cannot set reminders for the other circles). We refer to this as the "forced internal" condition.

They then saw a green button labelled, "Special circles worth 10 points. You must set reminders". Above this button was shown the instruction: "Other times, you will have to set reminders for all the special circles. When this happens, you will also score 10 points for every special circle you remember. In this case, you will be told 'This time you must set a reminder for every special circle' and you will see a green button. When this happens, the computer will make sure that you always set a reminder for every circle and it will not let you continue if you do not". After clicking on this they had one practice trial in which they had to set reminders. This was achieved by preventing participants from dragging circles out of the bottom of the box until after they had adjusted the location of forthcoming target circles.

Please note that while the description above implies that the red rectangle was always associated with the internal-memory option and the green rectangle with the external-reminder option, in fact the association of each strategy with red or green colour was randomised for each participant, as in Gilbert et al. (2020 experiment 3).

After participants practiced the two conditions, they were told, “Sometimes, you will have a choice between two options when you do the task. One option will be to do the task without being able to set any reminders. If you choose this option, you will always score 10 points for every special circle you remember. The other option will be to do the task with reminders, but in this case each special circle will be worth fewer points. For example, you might be told that if you want to use reminders, each special circle will be worth only 5 points. You should choose whichever option you think will score you the most points. So, if, for example, you thought you would earn more points by setting reminders and scoring 5 points for each special circle, you should choose this option. But if you thought you would score more points by just using your own memory and earning 10 points for each special circle you should choose this option instead”. Participants in the reward group were also reminded: “Please bear in mind that the more points you score, the more you will get paid at the end of the experiment”.

Experimental trials

Once participants were familiarised with how the scoring worked in the experiment, a series of 17 experimental trials began. Each of these 17 trials consisted of a sequence of 25 circles, including 10 targets. On even-numbered trials, participants performed either the forced internal or forced external condition in alternating order. The starting condition (either forced internal or forced external), was counterbalanced between participants. For the remaining 9 odd-numbered trials, they performed choice trials, where they could choose between an internal or external strategy. All possible target values when using reminders (from 1-9) were presented on these trials, in randomised order. Prior to these trials, participants were presented with both a red and a green rectangle, allowing them to choose between earning maximum points per target with an internal strategy (10 points) or a specified number of points (1-9) per target with an external strategy. These two options were presented next to each other with the left/right ordering of the internal/external options randomly counterbalanced between participants (see Gilbert et al., 2020).

Following each trial, participants in the no reward group were presented with the total number of points they had scored since the start of the experimental trials. Participants in the reward group were presented with the total number of points and money they had accumulated since the beginning of the experimental trials.

Second metacognitive judgement

After finishing the 17 experimental trials, participants were presented with the final metacognitive judgement scale which asked them to rate their accuracy on trials without reminders now that they had a full grasp of what the experiment entailed. The second metacognitive judgement scale was presented with the following set of instructions:

Now that you have had practice with the experiment, if you are presented with more trials, how accurately do you think you will be able to perform the task without any reminders? The difficulty of these trials would stay the same as the ones you have just completed. Please use *the scale below to indicate what percentage of the special circles you will be able to correctly drag to the instructed side of the square, on average, 100% would mean that you can always get every single one correct. 0% would mean that you can never get any of them correct.*

Participant inclusion criteria

The same inclusion criteria used by Gilbert et al. (2020 experiment 3), were used in this experiment. Participants were recruited from Amazon Mechanical Turk. Only participants based in the USA were included, to reduce variability in the sample. Furthermore, inclusion was restricted to participants with a minimum of 90% Mechanical Turk approval rate.

Participant exclusion criteria

The same exclusion criteria used by Gilbert et al. (2020 experiment 3), were used in this experiment. Participants were excluded if they satisfied any of the following criteria: a) accuracy in the forced internal condition equal or greater than accuracy in the forced external condition (which would imply that reminders do not improve performance, making data uninterpretable); b) accuracy in the forced internal condition and forced external condition lower than 10% and 70%, respectively; c) negative correlation between target value and the likelihood of choosing to use reminders, which would suggest random or counter-rational strategy choice behaviour. Furthermore, participants were also excluded if their reminder or metacognitive bias score was more than 2.5 standard deviations from the mean of their group. These participants were considered outliers. If participants were excluded for any of these reasons, additional participants were recruited so that the final sample consisted of 104 participants in each condition (208 in total).

Data Analysis

Independent variables

They key independent variable in this experiment was the reward group, which was manipulated between subjects.

Dependent variables

With the exception of the second metacognitive judgement scale, the dependent variables of this experiment were the same as the ones in our previous experiment. They were as follows:

1. Forced internal accuracy (ACC_{FI}). This was the mean target accuracy (i.e. proportion of target circles correctly dragged to the instructed location) on forced internal trials.
2. Forced external accuracy (ACC_{FE}). This was the mean target accuracy on forced external trials.
3. Optimal indifference point (OIP). This was the target value offered with reminders at which an unbiased individual should be indifferent between the two options, based on the ACC_{FI} and ACC_{FE}. As in Gilbert et al. (2020) this was calculated as:

$$OIP = (10 \times ACC_{FI}) / ACC_{FE}$$

4. Actual indifference point (AIP). This was the estimated point at which participants were actually indifferent to the two strategy options. As in Gilbert et al. (2020) this was calculated by fitting a sigmoid curve to the strategy choices (0 = own memory; 1 = reminders) across the 9 target values (1-9), using the R package 'quickpsy' bounded to the range 1-9.
5. Reminder bias. This was defined as OIP – AIP, which yielded a positive value for a participant biased towards using more reminders than would be optimal.
6. First metacognitive confidence. This was the response made to the metacognitive accuracy prediction.
7. First metacognitive bias. This was the difference between metacognitive confidence and actual accuracy on forced internal trials. A positive number would indicate overconfidence.
8. Second metacognitive confidence. This was the response made to the second metacognitive accuracy prediction
9. Second metacognitive bias. This was the difference between the second metacognitive judgement and actual accuracy on forced internal trials. A positive number would indicate overconfidence

Results

See Table 1 for a summary of results. All our analyses were conducted using R (version 4.0.1). We first investigated whether we could replicate our earlier findings (Gilbert et al., 2020 experiment 3) from the easy/positive group. These one-tailed analyses were restricted to the no-reward group only, which corresponded to the conditions of the earlier experiment. Participants were overconfident when they made their first metacognitive judgement (t(103)

= 2.43, $p = .0085$, $d = .24$) and also used reminders more often than was optimal, i.e. they had a positive reminder bias ($t(103) = 9.63$, $p < 10\text{-}15$, $d = 0.94$). Therefore, the earlier findings were replicated, showing both overconfidence in internal memory abilities and excessive use of reminders. We also conducted the same analyses for the reward group, but this time used two-tailed tests seeing as results in either direction could be theoretically informative. As in the no-reward group, participants in the reward group were overconfident ($t(103) = 2.05$, $p = .04$, $d = .20$) and biased towards reminders ($t(103) = 6.44$, $p < 10\text{-}8$, $d = .63$).

[Table 1 about here]

We next conducted our key analysis, comparing the reminder bias between participants in the reward and no-reward groups. A one-tailed independent samples t-test showed that the reminder bias was significantly reduced in the reward compared with the no-reward group ($t(203.7) = -2.85$, $p = .0024$, $d = .40$). This suggests that participants' excessive use of external reminders can be explained, at least in part, by a preference to avoid cognitive effort. Consistent with this, the no-reward group not only showed a greater bias towards reminders (versus the optimal strategy choice), but also a higher overall likelihood of choosing external reminders versus internal memory. This was reflected in both the total number of reminders used ($t(205.83) = 3.70$, $p = .00027$, $d = .51$) and the Actual Indifference Point ($t(205.66 = 3.71$, $p < .0001$, $d = .51$). We also found that the total number of points scored was lower in the no-reward group ($t(198.71) = 2.50$, $p = .01$, $d = .35$).

[Figure 2 about here]

In order to investigate whether participants' confidence changed between the first and second ratings, we conducted a mixed ANOVA on the first and second confidence ratings with Judgement (first, second) as the repeated measures factor and Group (reward, no reward) as the between-subjects factor. The main effect of Judgement was not significant ($F(1,206) = 3.43$, $p = .065$, $\eta^2_p = .016$), nor was the main effect of Group ($F(1,206) = 1.02$, $p = .315$, $\eta^2_p = .005$) or the interaction ($F(1,206) = .04$, $p = .846$, $\eta^2_p < .001$). This did not support the hypothesis that excessive use of reminders might be caused by a fall in confidence following the first metacognitive judgement (i.e. underconfidence with practice). We also investigated the second metacognitive bias measure in a univariate ANOVA with between-subject factor Group (reward, no-reward). Consistent with the first metacognitive judgement, rather than showing underconfidence the intercept of this ANOVA indicated overconfidence (mean = 2.45), but this was not significant ($F(1,206) = 2.34$, $p = .128$, $\eta^2_p = .011$). Nor was the main effect of Group ($F(1,206) < .0001$, $p = .987$, $\eta^2_p < .0001$).

[Figure 3 about here]

Next we investigated whether performance on the forced internal and external trials differed between the two reward groups. This was done by conducting a mixed measures ANOVA on target accuracy with factors Condition (forced internal, forced external) and Group (reward, no reward). There was a significant effect of Condition ($F(1,206) = 739.53$, $p < 10\text{-}6$, $\eta^2_p = .782$) but the effect of Group was not significant ($F(1,206) = 1.06$, $p = .31$, $\eta^2_p = .005$), nor was the interaction effect ($F(1,206) = 2.59$, $p = .11$, $\eta^2_p = .01$). Therefore, although the groups differed in their strategy choices (as shown above), they did not differ in their accuracy when forced to use one or the other strategy.

Furthermore, we investigated whether the reminder bias was related to metacognitive bias. We performed a simple linear regression with reminder bias as the dependent variable, and two independent variables: group (reward = 1, no-reward = -1) and the first metacognitive bias score. There was a significant effect of group (β = -.50, SE = .17, t(205) = -2.86, p < .01) but not metacognitive bias (β = -.01, SE = .01, t(205) = -1.01, p = .31). We also investigated whether the relationship between reminder bias and metacognitive bias differed between the groups. These bias scores were not significantly correlated in either the reward(r(102) = -.0076, p = .94) or the no-reward group (r(102) = -.13, p = .18). When the two correlation coefficients were transformed to z scores using Fisher's transformation they were not significantly different from each other (z = -.87, p = .38). All results were similar (i.e. significant results remained significant, nonsignificant results remained nonsignificant) when the second metacognitive bias score was used instead of the first.

That is, while the association between group and reminder bias was significant (β = -.50, SE = .17, t(205) = -2.86, p < .01), the second metacognitive bias score, like the first, was not (β < -.01, SE = .01, t(205) = -1.14, p = .25). Reminder bias was not significantly correlated with metacognitive bias in the no reward group (r(102) = -.12, p = .23) or the reward group (r(102) = -.03, p = .75), and these correlation coefficients were not significantly different from each other (z = -.62, p = .53).

Follow-up analysis

The analyses above correspond to the planned analyses detailed in the stage-1 submission of this registered report. We additionally performed one follow-up analysis that was not part of our original plan. Results above suggested that the groups did not differ in their accuracy when they used one or other strategy, but they did differ in their strategy choices. We also found that participants in the reward group earned more points than those in the no-reward group. This suggests that the reward group earned more points as a result of their strategy choices rather than their accuracy when they performed with one or other strategy. To test this hypothesis, we separated the total number of points scored on the forced strategy and the strategy-choice trials. We then conducted a mixed ANOVA with factors Group (reward, no reward) and Trial Type (forced, choice). There were significant main effects of Group ($F(1,206) = 6.24$, $p = .013$, $\eta^2_p = .029$) and Trial Type ($F(1,206) = 140.12$, $p < 10^{-24}$, $\eta^2_p = .405$), qualified by a significant interaction ($F(1,206) = 15.87$, $p < .0001$, $\eta^2_p = .072$). This indicates that the reward manipulation selectively influenced trials where participants had a choice of strategy.

Discussion

In some circumstances, individuals can be overconfident in their internal memory abilities yet still show a bias towards external reminders rather than internal memory (Gilbert et al., 2020 experiment 3). This counterintuitive finding suggests that metacognitive judgements cannot explain participants' bias towards external reminders in full. We investigated the hypothesis that an additional factor contributing to the reminder bias is a preference to avoid the cognitive effort associated with remembering intentions internally. We hypothesised that providing a financial incentive should motivate participants to invest more cognitive effort, thereby reducing the reminder bias. Results supported this hypothesis.

A secondary aim of this experiment was to investigate whether participants might become underconfident with practice (Koriat et al., 2002). In this case, participants' initial overconfidence might be reversed by the end of the experiment, which could also explain a bias towards external reminders. However, this hypothesis was not supported. If anything, participants were overconfident at the end of the experiment and their metacognitive judgements did not differ significantly between the beginning and end.

Cognitive Effort

Although the reminder bias differed significantly between the reward and no-reward groups, both groups chose external reminders more often than would have been optimal. Thus, provision of a financial incentive reduced but did not eliminate the reminder bias. This suggests that participants have a bias against cognitive effort that is not fully compensated by the rewards on offer. Consistent with this, Westbrook et al. (2013) found that participants were willing to accept a financial penalty to perform a task that was less cognitively demanding. Other studies have shown that the effect of financial incentives on cognitive effort is dependent on the relative value of those incentives (Otto & Vassena, n.d.; Rangel & Clithero, 2012; Tversky & Simonson, 1993). Therefore, individuals' decisions whether to use internal cognitive processes versus external tools may depend in a graded manner on differences in cognitive effort between the strategies and the rewards on offer. It is also of course possible that individuals select strategies not only based on metacognitive judgements and evaluations of effort, but also based on one or more additional factors which were not investigated here.

On the forced trials, when participants had to use one or other strategy, there was no significant difference in accuracy between the reward and no-reward groups. However, when there was a choice of strategy the reward group earned significantly more points than the no-reward group. Overall, these results corroborate previous findings that participants are more likely to expend cognitive effort when there is a financial incentive (Aarts et al., 2010; Botvinick & Braver, 2015; Padmala & Pessoa, 2011). Furthermore, they suggest this effect of financial incentive on effort allocation may be particularly evident when participants are given a choice of strategy, rather than being forced to use a particular approach.

Metacognition

One surprising result from this study was that metacognitive evaluations were not significantly correlated with participants' bias towards, or away from, use of external reminders. This contrasts with previous studies where such correlations have repeatedly been observed (Boldt & Gilbert, 2019; Gilbert, 2015b; Hu et al., 2019), including in experiments using the same paradigm that was used here (Engeler & Gilbert, 2020; Gilbert et al., 2020 experiments 2 & 3; Kirk et al., 2020). Seeing as this was not a primary focus of the present study, and it rests on a null result, it would be premature to draw strong conclusions. However, one possible interpretation would be that our metacognitive interventions (i.e. the practice and feedback procedures, which were designed to influence metacognitive judgements) reduced the validity of those judgements for predicting individual differences in strategic behaviour.

Conclusions

Recent work has shown that individuals decide whether to use internal cognitive processes or external tools based on various factors such as memory load and task interruption (Gilbert, 2015a), metacognitive beliefs and evaluations (Boldt & Gilbert, 2019; Dunn et al., 2016; Dunn & Risko, 2016; Gilbert, 2015b; Gilbert et al., 2020; Hu et al., 2019), age (Gilbert, 2015a), participants' past history and previous experience with act of offloading (Scarampi & Gilbert, 2020), and objective ability level (Gilbert, 2015b). This study shows that the availability of financial reward is another factor that influences such decisions, putatively because participants are more willing to allocate cognitive effort when a monetary reward is attached. Given that metacognitive evaluations of effort play a role in strategy selection (Dunn et al., 2016), this suggests a potential intervention to influence individuals' use of cognitive tools: making individuals aware of effort savings associated with cognitive offloading could influence their use of such strategies.

References

Aarts, E., Roelofs, A., Franke, B., Rijpkema, M., Fernández, G., Helmich, R. C., & Cools, R. (2010). Striatal dopamine mediates the interface between motivational and cognitive control in humans: Evidence from genetic imaging. *Neuropsychopharmacology*, *35*(9), 1943–1951. https://doi.org/10.1038/npp.2010.68

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology and Behavior*, *81*(2), 179–209. https://doi.org/10.1016/j.physbeh.2004.02.004

Boldt, A., & Gilbert, S. . (2019). Confidence guides spontaneous cognitive offloading. *Cognitive Research: Principles and Implications*, 1–33. https://doi.org/https://doi.org/10.31234/osf.io/ct52k

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113. https://doi.org/https://doi.org/10.1146/annurev-psych-010814-015044

Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science*, *24*(4), 255–260. https://doi.org/10.1177/0963721415569569

Dunn, T. L., Lutes, D. J. C., & Risko, E. F. (2016). Metacognitive evaluation in the avoidance of demand. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1372–1387. https://doi.org/https://doi.org/10.1037/xhp0000236

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

Engeler, N. C., & Gilbert, S. J. (2020). *The effect of metacognitive training on confidence and strategic reminder setting*. https://doi.org/https://doi.org/10.31234/osf.io/3vyxn

Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1280–1286. https://doi.org/10.1098/rstb.2012.0021

Francis, G., & Thunell, E. (2019). Excess success in "Ray of hope: Hopelessness increases preferences for brighter lighting." *Collabra: Psychology*, *5*(1), 1–7. https://doi.org/10.1525/collabra.213

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Fröber, K., & Dreisbach, G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive, Affective and Behavioral Neuroscience*, *14*(2), 530–547. https://doi.org/10.3758/s13415-014-0259-x

Gilbert, S. J. (2015a). 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). 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., Bird, G., Frith, C. D., & Burgess, P. W. (2012). Does "task difficulty" explain "task-induced deactivation?" *Frontiers in Psychology*, *3*(125), 1–12. https://doi.org/10.3389/fpsyg.2012.00125

Hu, X., Luo, L., & Fleming, S. M. (2019). A role for metamemory in cognitive offloading. *Cognition*, *193*(June), 104012. https://doi.org/10.1016/j.cognition.2019.104012

Jimura, K., Locke, H. S., & Braver, T. S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proceedings of the National Academy of Sciences*, *107*(19), 8871–8876. https://doi.org/10.1073/pnas.1002007107

Kirk, P. A., Robinson, O. J., & Gilbert, S. J. (2020). *Trait anxiety does not correlate with metacognitive confidence or reminder usage in a delayed intentions task*. https://doi.org/https://doi.org/10.31234/osf.io/m6twq

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*(4), 665–682. https://doi.org/10.2996/kmj/1138846322

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, *131*(2), 147–162. https://doi.org/10.1037/0096-3445.131.2.147

Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the stroop task. *Cognition*, *117*(3), 341–347. https://doi.org/10.1016/j.cognition.2010.08.018

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology*, *7*, 67–70. https://doi.org/10.1016/j.copsyc.2015.08.003

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Locke, H. S., & Braver, T. S. (2008). Motivational influences on cognitive control: Behavior, brain activation, and individual differences. *Cognitive, Affective and Behavioral Neuroscience*, *8*(1), 99–112. https://doi.org/10.3758/CABN.8.1.99

Niv, Y., Joel, D., & Dayan, P. (2006). A normative perspective on motivation. *Trends in Cognitive Sciences*, *10*(8), 375–381. https://doi.org/10.1016/j.tics.2006.06.010

Otto, A. R., & Vassena, E. (n.d.). It's all relative: Reward-induced cognitive control modulation depends on context. *Journal of Experimental Psychology: General*.

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, *23*(11), 3419–3432. https://doi.org/10.1162/jocn_a_00011

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: Theory and evidence. *Current Opinion in Neurobiology*, *22*(6), 970–981. https://doi.org/10.1016/j.conb.2012.07.011

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

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

Saunders, B., Lin, H., Milyavskaya, M., & Inzlicht, M. (2017). The emotive nature of conflict monitoring in the medial prefrontal cortex. *International Journal of Psychophysiology*, *119*, 31–40. https://doi.org/10.1016/j.ijpsycho.2017.01.004

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

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-

031526
Tversky, A., & Simonson, I. (1993). *Context- dependent Preferences*. *39*(10), 1179–1189. https://doi.org/https://doi.org/10.1287/mnsc.39.10.1179
West, J. T., & Mulligan, N. W. (2019). Prospective metamemory, like retrospective metamemory, exhibits underconfidence with practice. *Journal of Experimental Psychology: Learning Memory and Cognition*, *Advance online publication*.
Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? load, trait, and aging effects revealed by economic preference. *PLoS ONE*, *8*(7), e68210. https://doi.org/10.1371/journal.pone.0068210
Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

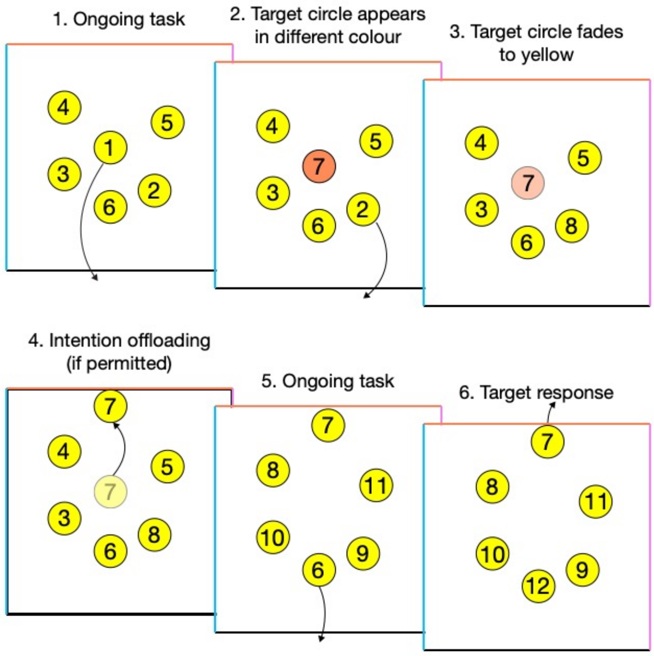

Figure 1. Schematic representation of the optimal reminders task

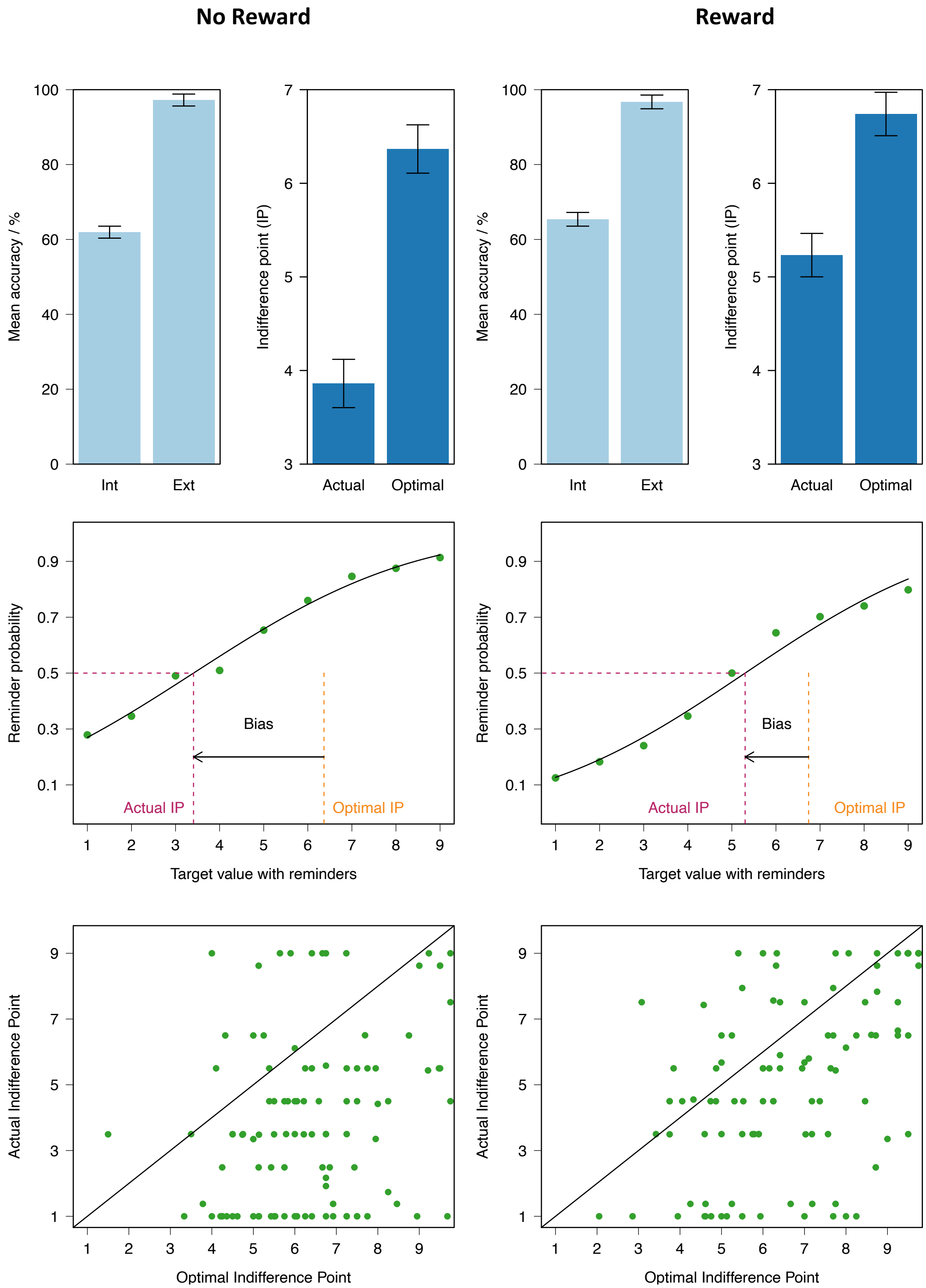

Figure 2. Behavioural results from experiment. Data from the no reward group is shown on the left and the reward group on the right. Top two graphs show mean accuracy in forced internal and forced external conditions, along with actual and optimal indifference points. Error bars represent within-subjects confidence intervals such that nonoverlapping bars indicate p < .05. The middle graphs depict the likelihood of

participants choosing to use reminders in choice trials (when a value between 1 and 9 was attached to this choice). Mean indifference points based on this graph are also shown. The graphs at the bottom illustrate each participants' optimal and actual indifference point. The diagonal line represents perfect calibration between the indifference points (i.e. actual = optimal). Points below this line indicate excessive use of reminders (actual < optimal) while points above this line indicate inadequate use of reminders (actual > optimal).

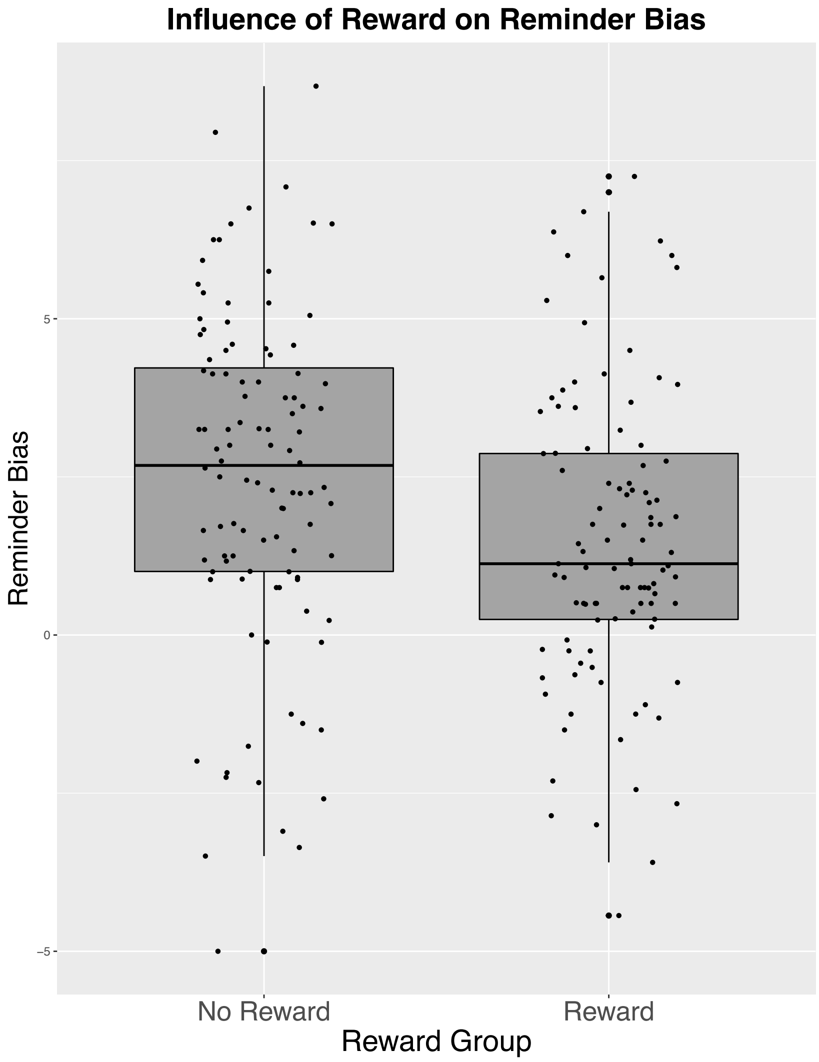

Figure 3. Boxplot depicting influence of reward on reminder bias. The two boxes illustrate the interquartile range of the data, along with the median. Minimum and maximum data points are shown through the bars. Individual observations have been added as jitter points to avoid dot overlap.

	No Reward	**Reward**
Forced external accuracy (%)	97.24 (4.86)	96.73 (5.19)
Forced internal accuracy (%)	61.95 (16.87)	65.38 (20)
Confidence Rating 1	67.72 (24.01)	70.48 (24.73)
Confidence Rating 2	64.42 (25.42)	67.81 (25.29)
Metacognitive Bias 1	5.77 (24.26)	5.10 (25.34)
Metacognitive Bias 2	2.48 (24.16)	2.42 (21.99)
AIP	3.86 (2.61)	5.23 (2.72)
OIP	6.37 (1.69)	6.74 (1.96)
Total reminders used	5.67(2.75)	4.27(2.67)
Total Points	1179.63(183.56)	1250.35(222.82)
Forced trial points	636.73(74.57)	648.46(89.32)
Choice trial points	542.90(122.90)	601.88(142.63)

Table 1. Behavioural results from both groups. Table shows means with standard deviations in parentheses. OIP = optimal indifference point; AIP = actual indifference point.