Title
The Impact of Meta-memory Judgments on Undergraduate’s Learning andMemory Performance.

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors
Humsani, Salwa Ali H
Civile, Ciro
McLaren, I. P.L.

Publication Date
2019

Peer reviewed

The Impact of Meta-memory Judgments on Undergraduate's Learning and Memory Performance.

Salwa Ali H Humsani (sh787@exeter.ac.uk)

Ciro Civile (c.civile@exeter.ac.uk)

I. P.L. McLaren (I.P.L.McLaren@exeter.ac.uk)

School of Psychology, University of Exeter, UK.

Abstract

We examined if using meta-memory judgments to control restudy choices has a positive impact on undergraduate students' memory performance, or whether simply making meta-memory judgments improved memory performance. 72 undergraduates at the University of Exeter were randomly divided into three groups. Participants in group A, had a chance to make meta-memory judgments and restudied the words they chose (self- selection). Participants in group B, also made meta-memory judgments, but restudy for this group was matched to that of Group A (control 1). Group C did not have a chance to make meta-memory judgments and were also matched to Group A for restudy opportunities (control 2). The results indicated that making meta-memory judgments had a positive overall impact on memory performance if undergraduates were allowed to control their restudy opportunities. Groups B and C showed no differences in memory performance, which means that making meta-memory judgments did not automatically improve memory performance. Group A restudied more of the words that they had rated as least well learned, and there were no significant differences between groups on test for the restudied words.

Keywords: Meta-memory Judgment (MJ), Restudy Choices, Learning, Memory.

Introduction

Meta-memory judgments rely on an individual's knowledge about how her or his memory processes affect their memory performance (Flavell, 1999; Hanczakowski, Zawadzka, & Cockcroft-McKay, 2014; Nelson, Dunlosky, Graf, & Narens, 1994). Efficient learning not only requires one to recognize information from memory, but also to be able to judge their level of confidence in material that they have previously learned and studied (Nelson, 1990; Nelson et al., 1994). One of the key reasons for studying meta-memory judgments is because it serves two functions: monitoring of memory processes and control over study behaviour (Nelson, 1990). The relationship between these functions is direct: people use memory monitoring, especially metamemory judgments, to decide which items need to be restudied and the length of time to be spent on them (Dunlosky & Hertzog, 1997; Kornell & Metcalfe, 2006). The central question addressed here is, do meta-memory judgments lead to effective study decisions? To test this experimentally, this study assumed that there would be positive effects on memory performance when participants are able to monitor their learning and are also able to use it to control their restudy opportunities.

Monitoring Accuracy: Studying cue-target word pairs is the most common approach used to investigate monitoring accuracy (Kimball, Smith, & Muntean, 2012; Nelson et al., 1994; Pyc, Rawson, & Aschenbrenner, 2014; Robey, Dougherty, & Buttaccio, 2017; Thiede & Dunlosky, 1994). This study design typically involves participants studying the word pairs, then making a Meta-memory judgment (MJ) to rate their ability to recall the target word when a cue is presented in the final test. Finally, participants take recall and recognition tests, which allow the researchers to assess how MJ can predict memory performance (Hughes, Taylor, & Thomas, 2018). Meta-memory judgments can be made immediately after the word pairs are studied, or delayed and made to the cue word alone. According to Nelson and Dunlosky (1991) the most important difference between immediate and delayed MJ lies in the amount of information available to participants when they judge their level of confidence. Participants who have made an immediate MJ have their target information in working memory, by contrast, this target is not available in working memory for a delayed MJ. Participants instead need to retrieve it from long-term memory. Several studies have investigated whether immediate MJ is more accurate than delayed MJ (Kimball et al., 2012; Nelson & Dunlosky, 1991; Pyc et al., 2014; Robey et al., 2017; Thiede & Dunlosky, 1994). While this study does not intend to assess the accuracy difference between these two processes, it uses the immediate MJ. The reason for choosing immediate MJ is provided by Hughes et al. (2018) who found that monitoring accuracy increases with immediate MJ when participants review material as a means of controlling repeated study or study- test practice. This study assumes that better meta-memory monitoring will lead to better restudy decisions, and also assumes that our main interest is in controlling restudy decisions at the time of study, rather than during later revision of the material.

Effectiveness of self-regulation: Effective learning involves two skills as stated earlier: monitoring learning and controlling study based on that monitoring (Kornell & Metcalfe, 2006; Nelson, 1990). Giving participants the opportunity to have control over their choices of which words to restudy allows them to be more engaged with their learning and improves their performance in the final memory tests. Begg, Martin & Needham (1992) and Hager & Hasselhorn (1992) concluded that self-memory monitoring is of no value to memory performance if participants did not control their study as well. In addition, Kornell & Metcalfe (2006) and Tullis & Benjamin (2012) tested the effectiveness of self-selection on using metacognitive judgements to control learning and memory performance. They found that allowing participants to control their learning had a positive effect on memory performance, long term learning and restudy choices. Methodologically, these researchers have used different ways to test the effectiveness of self-regulation on memory performance. Some of these ways involve comparing memory performance between groups: an experimental group (allowed to choose) versus a control group (choices made for them) (e.g. Begg et al., 1992; Kimball et al., 2012; Kornell & Metcalfe, 2006), as well as establishing comparisons on the basis of the best learned or worst learned restudied items (Nelson et al., 1994), or items rated as most difficult by participants (Thiede & Dunlosky, 1994). In this study, memory performance will be compared between all items, including those selected for restudy **or** unselected items across groups (experimental and control; Begg et al., 1992; Kimball et al., 2012; Kornell & Metcalfe, 2006). Several studies have shown that when allowing participants to judge their confidence and use that judgment to control their restudy decisions, the final memory performance was better than controls (Begg et al., 1992; Kimball et al., 2012; Kornell & Metcalfe, 2006; Nelson et al., 1994; Tullis & Benjamin, 2012). The first aim of this research is to test if using meta-memory judgment to control restudy choices has a positive impact on undergraduate students' memory performance, or whether simply making meta-memory judgments improved memory performance.

Method

Summary of the task: Forty concrete Arabic nouns with their pronunciations and translations were used to create two lists of word pairs, twenty in each list. Words were limited to be between four to eight letters. The words were randomly selected to serve as practice (first list) and study (second list) word pairs. Each phase of the experiment (practice followed by main study) started with instructions, then each item from the appropriate word list was presented on the screen for ten seconds on a white background. After each word, participants in Groups A (self-selection) and B had to judge their confidence of remembering the word in the future by rating their confidence from 1 to 9 (1=low confidence, 9= high confidence). Participants in Group C had to make a rating of how similar the Arabic word was to its English translation. Participants in Group A were then also asked if they needed to restudy the word just seen or not. Recall and recognition tests were given at the end of experiment.

Experiment design: The experiment used a between-subject design with three groups: Group A, who both made meta-memory judgments and could choose whether or not to restudy words (self-selection), control Group B, who made meta-memory judgments of their learning and experienced the same restudy opportunities as their counterpart in Group A (they were yoked to them), and control Group C, who did not make meta-memory judgments ratings but made similarity ratings instead, and were also yoked (in terms of restudy) to their counterpart in Group A. After seeing a word pair, participants first made their MJ to rate their meta-memory confidence of remembering the word later. Then, in Group A, each participant was able to request a restudy opportunity for any word. If they did, they then also made a meta-memory judgment after restudy as well. In Group B, each participant made a MJ, then got the same restudy opportunities as one of the participants in Group A but had no choice in the matter. If they were given a restudy opportunity, then they also made a second MJ to that word pair. For Group C, each participant had a chance to study the Arabic words, however they did not make a MJ during the task, but did get the restudy opportunities of one of the participants in Group A, and made a second similarity judgement after any restudy opportunity.

Participants: In this pilot study, random sampling was used. The participants of the study were 72 undergraduates from the University of Exeter; 24 in the meta-memory judgments and re-study (experimental group A), 24 in the MJ (control group B) and 24 in the No MJ (control group C) were random selected. The sample of the current study included 54 female and 18 male participants aged between 18 to 35 years, who did not speak the Arabic language and were enrolled in a variety of different subject disciplines. They were recruited via posters, email advertising and through the University of Exeter's Psychology Research Participation System. Participants were rewarded with a single payment of £5 or one credit on completion of the experiment, questionnaires and interview.

Procedure: All stimuli were presented with Superlab on a PC. The outline of the procedure for the study is summarised diagrammatically in Figure 1.

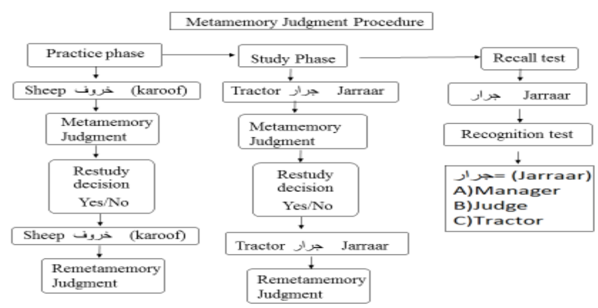

Figure 1 visual representation of the procedure.

Participants first completed a practice phase, which included 20 Arabic words with their translations and pronunciations in English. Then, participants in groups A and B made a judgment of how confident they were in their learning (selected number between 1 to 9). Whereas participants in group C made a judgment of how similar the English translation was to the Arabic pronunciation (again, a 1 to 9 rating, the purpose of this instruction was to have the same procedure across groups). After this, participants in group A were asked whether they would restudy this word if given the opportunity. Participants in group B and C were told that there was chance of repeating some words. Participants in groups A and B again made a meta-memory judgment after restudying words, whereas participants in group C again rated the similarity of the English translation word to the Arabic pronunciation. The same procedure was repeated in the study phase, but this included 20 new Arabic words. In the last phase, participants were asked to recall all the English translations cued by the Arabic words from the study phase in a random order. After this there was a recognition test provided for each word. One Arabic word with three English translations appeared on the screen, one of which was the right translation; participants had to choose that one. The following sections now give details of our procedure.

In the practice phase. Twenty Arabic words were presented in the middle of the screen between their translation and their pronunciation in English, with a font size of 16. A fixation stimulus was presented before the word pair and a blank screen after the word pair for 1 second. All participants studied the same word pairs (but in a random order), and all the word pairs were presented for ten seconds.

Meta-memory judgment: Participants in the meta-memory judgment A and B groups responded to the following instruction "Please select your level of confidence that you can remember this word pair by pressing the appropriate key from 1 to 9". Participants in group C in the non-meta-memory judgment did not make meta-memory judgments, but were asked the following question "Please rate the similarity of the English translation of this Arabic word to its Arabic pronunciation by pressing a key from 1 (very low) to 9 (very high)".

Restudy judgment. After making their meta-memory judgment, Participants in group A were asked "Would you like to re-study that Arabic to English translation? Press "y" for yes, "n" for no. If you've already re-studied once, then pressing either key will move you to the next trial". Participants in group B and group C were told "Note that there is a chance that the word pair you have just studied will be repeated. Press "y" to move on to the next trial". We arranged for participants in control group B and control group C to restudy the words determined by the matched participants in the experimental group A.

Second meta-memory judgment. After restudying a word, participants repeated the judgement appropriate for the group they were in.

Study Phase. In the main study phase, twenty new Arabic words with their translations and pronunciations in English were presented on the screen in a random order in a similar manner to the practice phase, and participants studied the pairs in order to remember them in the final recall test and recognise them in the final recognition test at the end of the experiment. All participants studied the same word pairs, and all the word pairs were presented for ten seconds at a font size of 16. They were then given meta-memory or similarity judgments and re-study opportunities as before.

Final test. After participants had completed the practice phase and study phase for all the 40 words, they completed a final recall test followed by a recognition test. More specifically, they were asked to recall the English translations of all the words from the study phase (i.e. all 20 words in that phase). They were given their Arabic form and pronunciations in English on the screen in a random order as a cue, and participants were asked to provide the English translation by typing their responses on the keyboard within 30 seconds. After all the words were tested in this way, a recognition test was given for each word. One Arabic word and English pronunciation with three English translations appeared on the screen, one of which was the right translation; participants had to choose this one. Another of the three translations was randomly taken from the practice phase, so that each practice word was used as a distractor once during this test phase. The other incorrect distractor word was novel, and would not have appeared before. None of the words used in a given test trial was repeated in any other test trial. Again, 30 seconds were given to do this, and once a word was selected they moved on to the next trial.

Results

74 participants were run on this experiment. Two participants were excluded because they did not complete the recall test. The results for the remaining 72 participants are as follows: The first issue we looked at was to determine if either of our control groups differed on either recall or recognition performance (the means for these groups are shown in Figure 2 below). They did not, both Fs<1, so we collapsed B and C into one overall control group and Compared this to A. An independent t- test used to examine the difference in the overall score between Experimental group A and this combined control group at recall gives a statistically significant difference between experimental group (M= 10, SD= 4.81) and control group (M=7.1, SD=3.78), t (70) = 2.727, $p<.001$, the eta squared statistic ($\eta2=.1$) indicates a large effect size A similar test used to examine the difference in the overall score for recognition also revealed a statistically significant difference, t = (70) = 2.558, p=.013. The effect size, calculated using eta squared, was close to large ($\eta2=0.09$). This implies that simply making a MJ does not automatically confer a significant benefit, but in combination with being able to choose which words to restudy it is effective in enhancing memory.

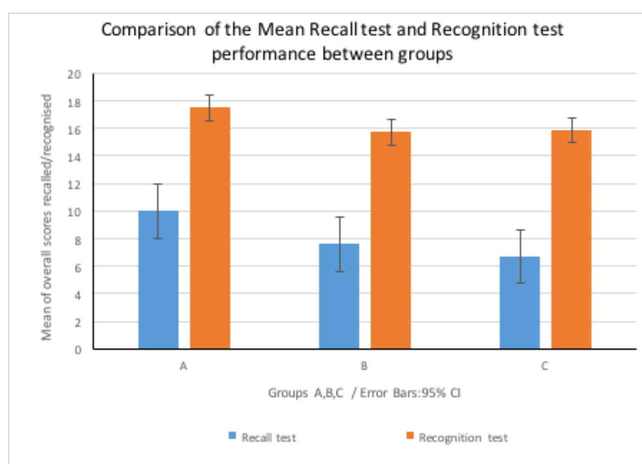

Fig. 2. This shows the difference between groups in their performance on the recall test and Recognition test.

The similar effects for recall and recognition suggest that both simply reflect memory for the word pairs, and this impression is confirmed by a correlational analysis. A Spearman rho test found a strong positive correlation at the 0.005 level (1-tailed) between the two variables, rho=.740.

Our next question was whether there was the expected relationship between the average MJ given to a word pair and performance on the recall and recognition tests? Obviously, we would expect higher average MJ to result in better performance on test. Correlations across participants failed to reveal any significant effects. When we compute this correlation across words, it failed to reveal a positive correlation between the MJ to a word pair and the recall test score for that word (r = .207, p=.079, n=48 1- tailed), but on the recognition test there was a significant and positive correlation between average MJ and performance (r= .439, p< .001, n=48, 1- tailed) with 19 % of the variance in recognition explained by this judgement. These results mean we have some evidence that some words are easier to learn than others.

In addition, as illustrated in Figure 3, there were differences in the MJ given before and after restudy for the restudied words, and also differences in the MJ given to non-restudied words. We analyzed this by performing two separate ANOVAs. The first was used to compare the MJ to the restudied and non-restudied words using the first judgement given in both cases (this would be the only judgement in the case of the non-restudied words). Group (A vs. B) was also included as a factor. The interaction between the groups factor (A, B) and the study factor (non-restudy, restudy) was significant, F (1, 46) = 7.100, p=.011. If we look at Figure 3, we can see that the MJs for Group A are higher than those for Group B for the non-restudied words, but lower for the words that were chosen by Group A for restudy. This is what is driving the interaction. An independent t- test showed that there was a statistically significant difference in the MJ between groups for the non- restudied words in favour of group A, t (46)= 3.073, p<.001, with large effect size (η^2=.3). The difference for the restudied words is not significant. It would appear, then, that Group A selected words that they found particularly difficult for restudy, leaving the easier words, and that this selection was somewhat specific to them, even though the Group B participants obviously show considerable agreement in what are the easier and harder words. This last point is reinforced by the main effect of study in this analysis, with the MJ for non-restudied words being much higher than that for the restudied words, F (1, 46) =40.69, p<.001.

The second ANOVA that we ran compared MJs to the restudied words before and after restudy. The interaction between groups (A, B) and study factor (before restudy and after) was just significant, F (1, 46) = 4.415, p=.041. Obviously, the effect of the factor of Study is stronger than this, the change from before to after is clear in Figure 3. The main effect for the type of MJ (before and after restudy) gave an F (1, 46) =35.197, p<.001, indicating that all participants show improved levels of confidence after restudying words. The interaction suggests that Group A improved more than Group B. Therefore, we have some evidence that restudy really helps participants to improve their level of confidence to remember the words in the final tests, and that effect was greatest when allowing participants to control their restudy opportunity.

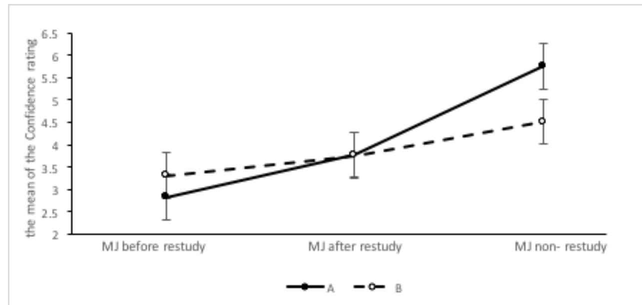

Fig 3. MJ before and after restudy and MJ for non-restudied words for groups A and B, Error Bars: 95%CI.

Turning now to a correlational analysis across participants, a Spearman rho test revealed a strong negative correlation between the mean of the initial MJ made by a participant and the number of requests for restudy in Group A, rho=.-.747 n=24, p<.001 (1-tailed). This correlation means that people who tended to give a lower MJ on average also tended to ask to restudy more. In essence, it could be taken to suggest that people have some sense of whether they are finding the task easy or hard, and adjust their study strategy accordingly. Another Spearman rho test gives a statistically significant and large positive correlation between frequency of restudy for a participant and their recall performance. That is, the more times people restudy (on average) the better their ability to retrieve words in the recall test, rho= .661 n=72, p<.001. There is a similar effect with recognition, a significant large positive correlation between frequency of study and recognition test performance, rho=.610, n=72, p<.001. It is not hard to see why this would be the case. The more restudy, the more practice of the items one gets, and if that helps then the better he or she performs. But this then leaves us with a slightly paradoxical situation, where the participants that we would argue are finding the task hardest (as signaled by a low MJ on average) are actually the ones performing the best. Perhaps the low average MJs may actually reflect better self-knowledge (i.e. a form of meta-memory) rather than ability as such. These are the people who know that they need to restudy, and do so, and benefit from it. Those with higher average MJs may be confident but may be mistaken in their confidence. It's also worth pointing out that the correlation between restudy request frequency and performance includes participants in Groups B and C who had no control over restudy. In some sense, the restudy manipulation was simply one imposed on them, and the result that more restudy benefitted performance is not surprising in that context. Finally, the correlations between MJ and test performance across subjects were not significant, so actually the paradox is not present in our data, just a potential feature of our theory.

After looking at the MJ data for the restudied and non-restudied words, we quite naturally would like to know how performance differed for those word types, and whether it differed across groups. Having collapsed B and C into one group as they do not differ on these measures, there was a main effect on recall for the type of words *F* (1, 69) =20.284, p=<.001. Participants performed better on non-restudied words than on restudied words. An independent *t*- test found statistically significant differences between the experimental group A (M= .504, SD= .240) and the combined control group (M=.360, SD=.192) for non- restudied words, *t* (70) = 2.734, p= <.001. The eta squared statistic (η^2= .1) indicated a large effect size. Whereas the difference between experimental group A (M= .301, SD= .368) and the control group (M=.222, SD=.315) for restudied words was not significant *t* (70) = .940, p=.350. The fact that overall, participants recalled more words when they were not restudied we take to simply reflect the fact that these were the easier word pairs. The fact that Group A was better than the combined control on these words again suggests an item specific advantage based on Group A participants selecting the words. Whilst the words were generally the easier ones (hence the main effect), the agreement on this between Group A and controls was not complete.

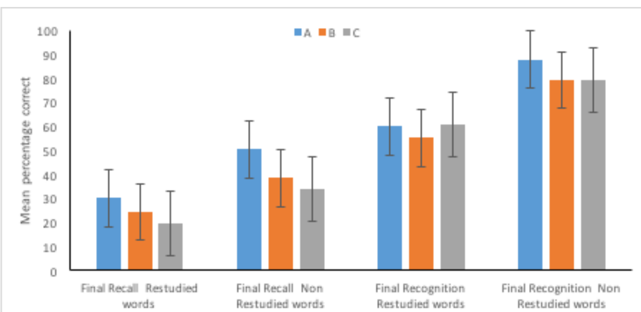

Fig. 4. Mean percentage of Final Recall and Recognition tests between restudied and non-restudied between groups A, B and C. Error Bars: 95%CI.

Turning now to the recognition test, once again participants recognise more words from the non-restudied words than the restudied words, *F* (1, 69) =20.783, p<.001. After collapsing B and C into one control group, an independent t- test showed there was a statistically significant difference between the experimental group (M= .878, SD= .111) and control group (M=.791, SD=.148) for the non- restudied words, *t* (70) = 2.515, p= =.014. The effect size, calculated using eta squared, was large (η^2= .1). However, the difference between the experimental group (M= .599, SD= .463) and control group (M=.579, SD=.423) for restudied words was not significant *t* (70) = .184, p=.854. (See Figure 4). Just as before, we can attribute some of this advantage on test to Group A's effective selection of the non-restudied words as the easier items, and that selection not transferring completely to the other groups. As a result, Groups B and C find these items harder on average and score lower on test. But the overall averages indicate general agreement about which are the more difficult items, as the overall means for the non-restudied words are much higher than for the restudied ones.

Discussion

In this experiment, we have found a significant advantage for Group A over the other two groups in terms of performance on the tests used to assess memory, with no significant differences between Groups B and C. It can be concluded that using meta-memory judgments and allowing control of restudy has a positive impact on participant's memory performance. The question now is why does this happen? Is it that the use of meta-memory allows our participants to select items that they find particularly difficult for restudy, thus improving performance? Or is it a more general effect? We can envisage two possibilities here. One is that people know how good their memories are and can make use of that knowledge to guide restudy. Another is that that giving people control over their choice of restudy items improves their motivation and engagement with the task. As we have seen and will see, there is evidence for both explanations, but what we can say is that simply "exercising" meta-memory, by giving a MJ to an item, is not in itself a significant factor in improving performance, otherwise there would be a significant difference between Groups B and C.

These results are in line with those of previous studies Begg, Martin, & Needham (1992) and Hager & Hasselhorn (1992) who concluded that self-memory monitoring has no value for memory performance if participants did not also have control of their study. In addition, these results are in agreement with those obtained by Kornell & Metcalfe (2006) and Jonathan G. Tullis & Benjamin (2012) who tested the effectiveness of self-selection on the use of metacognitive control over learning and memory performance. They found that allowing participants to control their learning had a positive effect on memory performance, long term learning and restudy choices.

This study found that there was a significant difference on MJ between non-restudy words and restudy words, and one possible explanation for this is that meta-memory monitoring helped participants in groups A and B to make their meta-memory judgments so as to discriminate between items which were more difficult and items that were easy and really sufficiently learned. These result are in line with those of previous studies who found that young subjects use their metacognition monitoring to distinguish between more difficult items (Li et al., 2018; Tullis & Benjamin, 2012; Tullis, Fiechter, & Benjamin, 2018; Zawadzka et al., 2018). Another significant finding was that meta-memory judgments could, to some extent, predict a participant's restudy frequency. Participants in group A, requested restudy more often for items that they had judged as least well – learned. This finding supports the work of other studies in this area (e.g. Dunlosky & Hertzog, 1997; Li et al., 2018; Nelson et al., 1994; Jonathan G. Tullis & Benjamin, 2012). Equally, participants in group A and B show significant improvement in their MJ after restudying words. These results agree with the finding of other study such as

Zawadzka et al. (2018) who demonstrate that repeated learning in the same environment improved learning, and metacognition monitoring. Obviously we are unable to comment on the effect of restudying a word on learning here because we do not know what performance on the restudied words would have been if they had not been restudied, but the positive correlation between frequency of restudying and test performance does fit in with the results cited. We intend to gather data that bears directly on this issue.

We can interpret some of the correlational results very simply as meaning that higher confidence about learning translates into better memory performance later. This would fit well with an item-specific effect of meta-memory on these tasks, whereby items judged as hard by Group A participants were given a low MJ, and this was used to trigger a restudy request. The effect of this was to improve performance on these items, back up to the level shown by the controls, while the advantage on the items not chosen for restudy because they were easy was greater in Group A again because they were able to make the right choices for them. Whilst there is general agreement about which are the easy and hard word pairs across groups (as shown by the correlations by word for MJ and test performance reported earlier) there is enough disagreement for the control groups to not gain as much benefit from the restudy offered, and so Group A does better. This is one possible explanation for our results.

But there may be more to this. Note that the MJ was, on average, higher in Group A than Group B, and, as we have seen, Group A performs better on test than Group B and C combined. The higher MJ in Group A could reflect increased confidence due to having control over which items are restudied, but this could be a general motivational effect rather than one based on meta-memory. To be clear, it could be that both the high MJ and better memory in Group A are both due to increased motivation due to their being in control of their restudy choices, in which case it would not be correct to say that the high MJ had some causal role in improving performance for Group A relative to Group B. Further research will be needed to disentangle the relationship between these variables.

Acknowledgments

This research is supported by The Saudi Arabian Cultural Bureau.

References

Begg, I. M., Martin, L. A., & Needham, D. R. (1992). Memory monitoring: How useful is self-knowledge about memory? *European Journal of Cognitive Psychology, 4*(3), 195-218.

Dunlosky, J., & Hertzog, C. (1997). Older and younger adults use a functionally identical algorithm to select items for restudy during multitrial learning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 52*(4), P178-P186.

Finn, B., & Metcalfe, J. (2008). Judgments of learning are influenced by memory for past test. *Journal of Memory and Language, 58*(1), 19-34. doi:https://doi.org/10.1016/j.jml.2007.03.006

Hager, W., & Hasselhorn, M. (1992). Memory monitoring and memory performance: Linked closely or loosely? *Psychological Research, 54*(2), 110-113.

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(3), 609.

Li, P., Zhang, Y., Li, W., & Li, X. (2018). Age-related differences in effectiveness of item restudy choices: the role of value. *Aging, Neuropsychology, and Cognition, 25*(1), 122-131.

Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. *Psychology of learning and motivation, 26*, 125-173.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect". *Psychological Science, 2*(4), 267-271.

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science, 5*(4), 207-213.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment Recall and Monitoring (PRAM). *Psychological methods, 9*(1), 53.

Souchay, C., & Isingrini, M. (2012). Are feeling-of-knowing and judgment-of-learning different? Evidence from older adults. *Acta Psychol (Amst), 139*(3), 458-464. doi:10.1016/j.actpsy.2012.01.007

Thiede, K. W., & Dunlosky, J. (1994). Delaying students' metacognitive monitoring improves their accuracy in predicting their recognition performance. *Journal of Educational Psychology, 86*(2), 290.

Tullis, J. G., & Benjamin, A. S. (2012). Consequences of restudy choices in younger and older learners. *Psychonomic Bulletin & Review, 19*(4), 743-749. doi:10.3758/s13423-012-0266-2

Tullis, J. G., Fiechter, J. L., & Benjamin, A. S. (2018). The efficacy of learners' testing choices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(4), 540.

Zawadzka, K., Simkiss, N., & Hanczakowski, M. (2018). Remind me of the context: Memory and metacognition at restudy. *Journal of Memory and Language, 101*, 1-17.