Routledge
Taylor & Francis Group

State-based metacognition: How time of day affects the accuracy of metamemory

Kathleen L. Hourihan[1] and Aaron S. Benjamin[2]

[1]Department of Psychology, Memorial University of Newfoundland, St John's, NL, Canada
[2]Department of Psychology, University of Illinois at Urbana-Champaign, IL, USA

Although there is an abundance of research on how stimulus characteristics and encoding conditions affect metamemory, and how those effects either do or do not mirror effects on memory, there is little research on whether and how characteristics of participants' states—like mood, fatigue, or hunger—affect metamemory. The present study examined whether metamemory ability fluctuates with time of day. Specifically, we evaluated whether learners can successfully account for the effects of time of day on their memory, and whether metacognitive monitoring is more accurate at an individual's optimal time of day. Young adults studied and recalled lists of words in both the morning and the afternoon, providing various metamemory judgements during each test session. We replicated the finding that young participants recalled more words in the afternoon than in the morning. Prior to study, participants did not predict superior recall in the afternoon, but they did after they had an opportunity to study the list (but before the test on that material). We also found that item-by-item predictions were more accurate in the afternoon, suggesting that self-regulated learning might benefit from being scheduled during times of day that accord with individuals' peak arousal.

Keywords: Metamemory; Time of day; Memory.

The study of *metamemory* assesses what we think about our own memory and its performance. Studying metamemory has important practical implications in many domains including, for example, ongoing self-assessments of learning for students preparing for an exam, decisions about how to improve performance in domains in which learning is below desired levels (e.g., Finley, Tullis, & Benjamin, 2009; Thiede & Dunlosky, 1994, 1999; Tullis & Benjamin, 2012), and assessments of confidence and accuracy in remembering the details of a crime (e.g., Leippe & Eisenstadt, 2007).

The accuracy of metamemory monitoring is usually evaluated by asking participants either to make a global assessment of their learning following a learning phase (but usually prior to test), or to make item-by-item judgements during the learning phase—*Judgements of Learning* (JOLs), (Arbuckle & Cuddy, 1969; Nelson, 1984). Accuracy in such judgements is revealed by a high degree of correspondence between judgements and later performance, assessed by either direct comparison of ratings with performance (calibration, or absolute accuracy) or the correlation, across items or groups of items, of

judgements and performance (resolution, or relative accuracy).

In this paper we extend the study of the sensitivity of JOLs to the understudied domain of *state-based* characteristics—that is, characteristics that participants bring with them to the laboratory. Specifically, we examine whether metamnemonic accuracy varies at different times of day. We chose this domain because actual memory performance is known to vary systematically with time of day (e.g., Anderson, Petros, Beckwith, Mitchell, & Fritz, 1991; May, Hasher, & Stoltzfus, 1993), and because a metacognitive appreciation of time-of-day effects would seem to have interesting ramifications for self-guided study.

Although it is generally well accepted that learners undervalue the importance of learning, orienting, and forgetting (Koriat, 1997; Kornell & Bjork, 2009) and overvalue the importance of intrinsic stimulus characteristics (Koriat, 1997), the applicability of this generality to state-based variables like time of day is unclear. Learning about the influence of time of day on memory requires access to compatible measures of memory across those different times, something that is—at least for the average student—usually unavailable. If calculus is in the morning and history in the afternoon, that student can never tease apart the effects of time of day from the effects of the material and course content.

There are a handful of interesting examples in the literature of state-based metamemory, but the methodologies and measures across those studies are variable, and the results do not paint a consistent picture. For example, one study (Nelson, McSpadden, Fromme, & Marlatt, 1986) found that alcohol impaired memory but did not affect confidence or feeling of knowing (FOK) accuracy (see also Darley, Tinklenberg, Roth, Vernon, & Kopell, 1977). Other cases indicate that some state-based factors do affect metamemory. Nelson et al. (1990) examined how memory and metamemory were affected by hypoxia due to extreme altitudes. Although memory retrieval was unaffected by extreme altitudes, absolute metamemory accuracy (as measured by confidence and FOK judgements) was negatively affected. The effects of caffeine on metamemory (Kelemen & Creeley, 2003) revealed insensitivity of mean JOLs to state, but superior metamemory when study and test states were mismatched than when matched (see also Merritt, Hirshman, Hsu, & Berrigan, 2005; Mintzer & Griffiths, 2005).

All of the studies discussed so far that examined state-based effects on metamemory required some sort of external manipulation. That is, participants were given a dose of alcohol, caffeine, or other drugs, with the research goal of determining how these interventions affect memory and metamemory. But does metamemory ability fluctuate with internal factors that do not require intervention from an experimenter? It has been known for a long time that various cognitive abilities, including memory, may be affected by the fluctuations in subjective arousal that occur naturally throughout the day (e.g., Baddeley, Hatter, Scott, & Snashall, 1970; Blake, 1967; Folkard, Monk, Bradbury, & Rosenthal, 1977; May et al., 1993; Millar, Styles, & Wastell, 1980). The goal of the present study is to determine whether metacognitive judgements are sensitive to this effect, and also to evaluate whether metamemory ability fluctuates with time of day.

Metamemory ability is critical to study in an academic context because student learning is an active and self-regulated process: Metacognitive processes underlie decisions about organising and scheduling study of materials, as well as decisions about mastery of those materials (Benjamin, 2003; Finley et al., 2010). The present study examined whether the metamemory (and memory) ability of young adults is better at their optimal time of day—the afternoon—than at a suboptimal time of day—early morning. Our focus is on two specific questions. First, do JOLs reveal sensitivity to the effects of time of day? That is, do learners accord higher ratings to material learned at preferred times of day? Second, does the relative metamnemonic accuracy of those judgements differ across times of day? The former question is relevant to whether learners will be able to successfully account for the effects of time of day on their memory; the latter is relevant to whether self-regulated learning is likely to be more effective at one time of day or another.

METHOD

Participants

A total of 58 students at the University of Illinois at Urbana-Champaign participated in exchange for bonus course credit or payment.

Materials

The stimulus pool consisted of 144 words obtained from the online version of the MRC Psycholinguistic Database (Coltheart, 1981). All words were four to nine letters in length, and were relatively high in frequency (ranging from 7 to 2244, $M = 303$; Kučera & Francis, 1967). Words were randomly selected for study from the pool of 144 items on a per-participant basis using E-Prime software.

At the end of the second testing session all participants completed a standardised morningness/eveningness questionnaire (Horne & Östberg, 1977) to confirm whether our younger adult participants were in fact more likely to be evening-type individuals. This questionnaire is widely used in studies examining time of day effects and has been shown to correlate strongly with physiological measures of peak circadian arousal that differ across individuals.

Design and procedure

Each participant completed two sessions, one in the morning (either at 8 am or 9 am) and one in the afternoon (either at 3 pm or 4 pm); session order was counterbalanced. The second session took place 6–8 days following the first session. The two testing sessions were nearly identical, with the exception that participants completed questionnaires at the end of their second testing session. In each session participants completed two study–test cycles on a computer, seated alone in a small room. An instruction screen informed participants that they would be studying a list of 20 unrelated words that they would have to recall and were asked to make a global prediction of how many words, out of 20, they thought they would recall. Following the first prediction an instruction screen was displayed that described the JOL judgements. In the study phase one word was presented in the centre of the screen for 2000 ms, followed by a blank screen for 1000 ms. The JOL screen was then presented. On the JOL screen participants were instructed to press a key from 1 to 6 to indicate how likely they thought it was that they would remember the word just presented. A 1000-ms blank screen preceded the next trial. After all 20 words had been presented (and JOLs collected), participants were asked to make a second global prediction of how many

words out of 20 they thought they would recall. They then completed a distractor task consisting of 30 seconds of simple maths problems, and were informed that they would have 2 minutes to type in as many words as they could recall, in any order.

Following free recall, participants completed another minute of maths problems. The second study–test cycle then began, and was nearly identical to the first cycle except with new words randomly selected from the pool, and no final maths distractor task following recall. The second session was nearly identical to the first session, but with new words randomly selected from the stimulus pool for both cycles.

At the end of the second test in the second session all participants completed an open-ended questionnaire which asked participants whether they thought they had remembered more words in one test session or the other (the morning or the afternoon), as well as whether they thought that their memory in general was affected by time of day. Participants then completed the standardised morningness/eveningness questionnaire (Horne & Östberg, 1977).

RESULTS AND DISCUSSION

List number within a session did not yield any significant effects, and all means reported below are averaged across both lists. We also did not observe any significant effects of session order, nor did order interact with any other factors. We therefore collapsed across this factor in all reported analyses.

Questionnaires

Participants were classified into types based on their score on the morningness/eveningness questionnaire (Horne & Östberg, 1977). The majority of the 58 participants were classified as either "neither type" ($n = 31$) or as "moderately evening types" ($n = 20$). Four participants were classified as "definitely evening types" and the remaining three participants were "moderately morning types" (no participants were classified as "definitely morning types"). In the analyses below, no effects of morningness/eveningness type or score were found (likely because the majority of participants are classified into two adjacent categories). We therefore included all participants and

do not include the questionnaire score or type as a factor.

On the open-ended questionnaire the majority of participants reported better memory in the afternoon, both for experimental performance (53%) and in general (43%). Fewer participants reported better memory in the morning (24% for experimental performance and 29% for memory in general). The remaining participants reported no systematic time of day differences in memory. Participants' intuitions were generally correct. Actual differences in memory performance between the two test sessions correlated significantly with participants' beliefs about the session in which they performed best ($r = .63$, $p < .001$), and the time of day in which they believed their memory was best in general ($r = .46$, $p < .01$).

Memory accuracy

Accuracy on the free recall tests is displayed in the first row of Table 1. A t-test showed that participants recalled more words in the afternoon than in the morning, $t(57) = 2.65$, $p < .05$, replicating the standard effect of time of day on memory accuracy in younger adults.

Metamemory judgements

Global judgements. Given our interest in participants' sensitivity to memory differences at the different times of day, we examined separately whether the first predictions (i.e., the predictions made prior to list study) and second predictions (i.e., the predictions made after list study but prior to recall) showed time of day effects. Means are displayed in Table 1. Results showed that predictions made prior to study were not sensitive to time of day, $t(57) = 1.05$, but predictions made after study and prior to test were significantly higher in the afternoon than in the morning, $t(57) = 2.79$, $p < .01$, accurately reflecting the superior level of recall observed in the afternoon.

Mean JOLs. The mean JOL ratings were computed across words in the morning and afternoon session, and are displayed in Table 1. Participants did not give higher or lower JOLs, on average, in either test session, $t(57) = 0.54$, $p = .59$.

Correspondence between judgements and performance. We computed each participant's correspondence between immediate JOLs and recall accuracy for the morning and afternoon sessions using both the more traditional Goodman-Kruskal gamma correlation, G (e.g., Nelson, 1984), and the signal detection-based d_a statistic (cf. Benjamin & Diaz, 2008; Masson & Rotello, 2009). The values for both measures are displayed in Table 1, but we focus here on the d_a statistic because of its desirable measurement properties (see Benjamin & Diaz, 2008; Masson & Rotello, 2009). Paralleling the results observed for memory accuracy, metamemory resolution was numerically better in the afternoon than in the morning, $t(57) = 1.90$, $p = .063$, approaching conventional levels of statistical significance. A sign test revealed that the majority of participants displayed superior resolution in the afternoon session ($Z = 1.97$, $p < .05$).

TABLE 1
Memory and metamemory measures for the morning and afternoon testing sessions

Measure	Time of day		pm − am difference
	Morning (am)	Afternoon (pm)	
Free Recall (# words)	9.63 (0.38)	10.44 (0.34)	+0.81 (0.31)
First Prediction (# words)	8.02 (0.34)	8.33 (0.34)	+0.30 (0.29)
Undentfidence	− 1.61 (0.30)	− 2.11 (0.38)	− 0.51 (0.43)
Second Prediction (# words)	7.94 (0.35)	8.65 (0.36)	+0.72 (0.26)
Undentfidence	− 1.69 (0.29)	− 1.78 (0.37)	− 0.10 (0.33)
Mean JOL (1 − 6)	3.44 (0.07)	3.49 (0.08)	+0.04 (0.07)
Metamemory Resolution			
Gamma (G)	.26 (.04)	.35 (.04)	+.09 (.06)
D_a	0.35 (0.06)	0.52 (0.06)	+0.17 (0.09)

Standard error is displayed in parentheses beside each mean.

GENERAL DISCUSSION

The goal of the present study was to determine the sensitivity of metacognitive judgements to a state-based factor: time of day. We replicated the finding that young participants recalled more words in the afternoon than in the morning. Importantly, we found evidence that participants were aware of the effects of time of day on memory; questionnaire results indicated that most participants believed their memories were better in the afternoon, and participants predicted recalling more words in the afternoon session than in the morning session. Intriguingly, participants' first predictions (made prior to study) did not reflect the time of day effect in memory, but their second predictions (made after study but prior to test) were significantly higher in the afternoon than in the morning. Participants seem not to have expected to remember a list of words better in the afternoon, prior to studying it. However, the experience of actually studying the list led to higher post-study recall predictions in the afternoon than in the morning, accurately reflecting the superior recall observed in the afternoon session.

One possibility to explain this finding is that participants experienced an unanticipated feeling of fluency during the process of encoding in the afternoon session. That is, although the predictions made prior to list study did not reflect time of day differences in memory, learning might have simply felt easier in the afternoon, leading participants to increase their recall predictions after having studied the list. Given that an individual's optimal time of day is related to subjective arousal and alertness, it may be that participants had to actually experience an attempt to learn to appreciate how their present state might affect their memory. This idea fits with our finding that, while only 43% of our participants believed that their memory in general is better in the afternoon, 53% of our participants recognised that their performance in the experiment was better in the afternoon: Experiencing the act of encoding appears to help inform metamemory judgements. Another possibility is that participants might have been covertly attempting to recall the list items when making the second prediction. Awareness of being able to recall more list items in the afternoon session would lead to the accurate (second) prediction of greater recall in the afternoon, relative to the morning.

The additional finding that item-by-item predictions are less accurate at a suboptimal time of day has significant importance for students. An important aspect of the learning processes involved in taking a college course is self-assessment: Students must learn to work independently and to prepare themselves sufficiently for exams with their own study strategies. To determine whether he or she has studied enough to score well on tomorrow's exam, a student must be able to evaluate whether they will be able to remember this information when they try to retrieve it in the future. The implication of our results is that most students will be able to answer this question more accurately in the afternoon than in the morning.

CONCLUSIONS

The goal of the present study was to determine whether (and how) metamnemonic accuracy might change as a function of time of day. We found that individuals have some understanding of the changes in memory from optimal to suboptimal time of day, but that understanding only reveals itself in global judgement made following a study opportunity. In addition, our young adult participants exhibited significantly poorer correspondence between judgements and performance in the morning than in the afternoon. Although studying is nearly always likely to be a worthwhile activity for students, our results suggest that students should be cautious about assessing their own learning when studying course material at times other than their own optimal time of day.

REFERENCES

Anderson, M. J., Petros, T. V., Beckwith, B. E., Mitchell, W. W., & Fritz, S. (1991). Individual differences in the effect of time of day on long-term memory access. *American Journal of Psychology*, *104*, 241–255.

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at the time of presentation. *Journal of Experimental Psychology*, *81*, 126–131.

Baddeley, A. D., Hatter, J. E., Scott, D., & Snashall, A. (1970). Memory and time of day. *Quarterly Journal of Experimental Psychology*, *22*, 605–609.

Benjamin, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition, 31*, 297–305.

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of memory and metamemory* (pp. 73–94). New York, NY: Psychology Press.

Blake, M. J. F. (1967). Time of day effects on performance in a range of tasks. *Psychonomic Science, 9*, 349–350.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.

Darley, C. F., Tinklenberg, J. R., Roth, W. T., Vernon, S., & Kopell, B. S. (1977). Marijuana effects on long-term memory assessment and retrieval. *Psychopharmacology, 52*, 239–241.

Finley, J. R., Tullis, J. G., & Benjamin, A. S. (2009). Metacognitive control of learning and remembering. In M. S. Khine & I. M. Saleh (Eds.), *New science of learning: Cognition, computers and collaboration in education*. New York, NY: Springer Science & Business Media.

Folkard, S., Monk, T. H., Bradbury, R., & Rosenthall, J. (1977). Time of day effects in school children's immediate and delayed recall of meaningful material. *British Journal of Psychology, 68*, 45–50.

Horne, J., & Östberg, O. (1977). A self-assessment questionnaire to determine morningness–eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*, 97–110.

Kelemen, W. L., & Creeley, C. E. (2003). State-dependent memory effects using caffeine and placebo do not extend to metamemory. *The Journal of General Psychology, 130*, 70–86.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgements of learning. *Journal of Experimental Psychology: General, 126*, 349–370.

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General, 138*, 449–468.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lieppe, M. R., & Eisenstadt, D. (2007). Eyewitness confidence and the confidence–accuracy relationship in memory for people. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology, volume II: Memory for people* (pp. 377–425). Mahwah, NJ: Erlbaum.

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory & Cognition, 35*, 509–527. doi:10.1037/a0014876

May, C. P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychological Science, 4*, 326–330.

Merritt, P., Hirshman, E., Hsu, J., & Berrigan, M. (2005). Metamemory without the memory: Are people aware of midazolam induced amnesia? *Psychopharmacology, 177*, 336–343.

Miller, K., Styles, B. C., & Wastell, D. G. (1980). Time of day and retrieval from long-term memory. *British Journal of Psychology, 71*, 407–414.

Mintzer, M. Z., & Griffiths, R. R. (2005). Drugs, memory, and metamemory: A dose–effect study with lorezepam and scopolamine. *Experimental and Clinical Psychopharmacology, 13*, 336–347.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*, 109–113.

Nelson, T. O., Dunlosky, J., White, D. M., Steinberg, J., Townes, B. D., & Anderson, D. (1990). Cognition and metacognition at extreme altitudes on Mount Everest. *Journal of Experimental Psychology: General, 119*, 367–374.

Nelson, T. O., McSpadden, M., Fromme, K., & Marlatt, G. A. (1986). Effects of alcohol intoxication on metamemory and on retrieval from long-term memory. *Journal of Experimental Psychology: General, 115*, 247–254.

Thiede, K. W., & Dunlosky, J. (1994). Delaying students' metacognitive monitoring improves their accuracy in predicting their recognition performance. *Journal of Educational Psychology, 86*, 290–302.

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1024–1037.

Tullis, J. G., & Benjamin, A. S. (2012). The effectiveness of updating metacognitive knowledge in the elderly: Evidence from metamnemonic judgements of word frequency. *Psychology and Aging, 27*, 683–690.