

Metacognition and the spacing effect: the role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal

Jessica M. Logan · Alan D. Castel · Sara Haber ·
Emily J. Viehman

Received: 10 April 2012 / Accepted: 17 September 2012 /
Published online: 26 September 2012
© Springer Science+Business Media New York 2012

Abstract Although memory performance benefits from the spacing of information at encoding, judgments of learning (JOLs) are often not sensitive to the benefits of spacing. The present research examines how practice, feedback, and instruction influence JOLs for spaced and massed items. In [Experiment 1](#), in which JOLs were made after the presentation of each item and participants were given multiple study-test cycles, JOLs were strongly influenced by the repetition of the items, but there was little difference in JOLs for massed versus spaced items. A similar effect was shown in [Experiments 2 and 3](#), in which participants scored their own recall performance and were given feedback, although participants did learn to assign higher JOLs to spaced items with task experience. In [Experiment 4](#), after participants were given direct instruction about the benefits of spacing, they showed a greater difference for JOLs of spaced vs massed items, but their JOLs still underestimated their recall for spaced items. Although spacing effects are very robust and have important implications for memory and education, people often underestimate the benefits of spaced repetition when learning, possibly due to the reliance on processing fluency during study and attending to repetition, and not taking into account the beneficial aspects of study schedule.

Keywords Metamemory · Memory · Spacing effects · Judgments of learning

Author Note We would like to thank David Balota, Aaron Benjamin, Bob Bjork, Elizabeth Bjork, Nate Kornell, David McCabe and Matthew Rhodes for valuable comments at various points during this project. Pooja Agarwal was extremely helpful in collecting and analyzing data and providing insight.

J. M. Logan (✉) · E. J. Viehman
Department of Psychology, Rice University, P.O. Box 1892, MS-25, Houston, TX 77251-1892, USA
e-mail: Jessica.Logan@rice.edu

A. D. Castel (✉)
Department of Psychology, University of California, Los Angeles, 1285 Franz Hall Box 951563, Los Angeles, CA 90095-1563, USA
e-mail: castel@ucla.edu

S. Haber
The Center for Vital Longevity, University of Texas, Dallas, TX, USA

Memory performance benefits from the repeated presentation of items, and long-term retention benefits when these items are spaced apart in time, rather than massed. This “spacing effect” has been demonstrated in a number of instances, with different populations, and is a highly robust phenomenon (see Cepeda et al. 2006, for a review; see also Glenberg 1976). On an applied level, spaced schedules would be ideal for students when studying for tests or exams. However, spaced strategies are often not employed, and several lines of evidence suggest that participants often fail to appreciate the benefits accrued by spaced practice. For example, Baddeley and Longman (1978) demonstrated that postal workers preferred massed over distributed training despite the fact that spaced practice lead to better performance. Zechmeister and Shaughnessy (1980) showed that participants’ memory predictions did not differentiate between massed and spaced rehearsal (or were more likely to favor massed rehearsal), despite the fact that recall was superior for spaced items. Evidence from motor learning also suggests that, although massed practice can benefit short-term retention, participants predict that massed practice will also lead to better later retention, relative to spaced practice (Simon and Bjork 2001). Thus, prior data suggests that, contrary to actual performance, participants consistently regard massed practice as better for learning than spaced practice (cf. Bjork 1999). The present study examines this issue by using a design that includes experience-based learning and feedback about recall performance.

Other methods of assessing metacognitive performance are consistent with the idea that participants do not fully appreciate the benefits of spacing. For example, several studies have allowed participants to choose to employ different spacing schedules. When items are presented at a fast presentation rate, Son (2004) showed that participants chose massed practice for certain items, relative to spacing. Thus, even when studying is under participants’ control, they choose massed practice. Benjamin and Bird (2006) reported that participants would space items when told that they needed to space half of the items (and mass the remaining items), such that participants will space more difficult items that they feel they need to practice at a later time (see also Toppino et al. 2009). Pyc and Dunlosky (2010) found that participants chose to mass both easier and rapidly-presented items. Toppino and Cohen (2010) also noted a preference among participants to space study of difficult items during a cued-recall task. They found that subjects chose to space more items that were given an arbitrary high “point value” (5 points vs. 1 point or no points) for memory, regardless of difficulty level. Taken together, they argue that these findings support Ariel et al. (2009) agenda-based regulation model, which states that learners create study agendas based on maximizing goal achievement.

These studies suggest that although participants may choose study options that include spacing items under certain circumstances, it is not clear if participants know that spacing will lead to better recall, and if predictions of later recall will reflect the benefits of spaced versus massed rehearsal. Individuals tend to greatly underestimate the effect spacing has on memory performance. This can lead to a general “illusion of comprehension” or competence (e.g., Jacoby et al. 1994), in which students feel better prepared with massed practice, despite spaced practice leading to better later memory performance. The current research extends upon previous work by exploring whether providing subjects with experience, feedback, and information on the benefits of spacing would help them become more aware of the spacing effect.

In the context of judgments of learning (JOLs) and spaced versus massed practice, participants often underestimate the influence of spacing on later recall. Zechmeister and Shaughnessy (1980) asked participants to make memory predictions after a single presentation of an item, or after the second presentation of items that were presented twice in either a massed or spaced manner. With a short list of words (24) and a delayed free recall task after a one-minute retention interval, participants’ predictions were higher for the twice versus once presented items, but participants were more inclined to give slightly higher JOLs for

massed than spaced items, despite a recall advantage for spaced items. This design may be improved by asking people to make predictions after each presentation of each item (as opposed to only after the second presentation), as this may lead to more accurate predictions.

Research in the area of metacognition suggests that individuals typically prefer massed practice (or fail to understand the benefits of spacing), but that this may be the result of participants having little experience with retrieval following massed or spaced practice. For example, Koriat et al. (2004) have shown that participants' JOLs are insensitive to retention interval, and that participants place high priority on the properties of the items (e.g., associative relatedness of word pairs), as opposed to test-related information such as the length of retention interval between study and test. Several studies (e.g., Dunlosky and Hertzog 2000; Koriat and Bjork 2006) have shown that practice with encoding and retrieval conditions can lead to alterations in predictions that better capture memory performance. For instance, deWinstanley and Bjork (2004) found that after studying passages in which both to-be-read and to-be-generated items were present and being tested, people would demonstrate a traditional generation advantage in memory. However, on a subsequent passage and test, participants improved their performance on to-be-read items to match that of the generated items, suggesting some use of prior experience informing later encoding strategies. In a similar vein, Castel (2008) found that participants learned to incorporate serial position information when making JOLs, but only with experience and when the serial position information was easily accessible during encoding.

Why do people's memory predictions not differentiate between massed and spaced rehearsal, despite memory being strongly affected by spaced practice? In the present study, we examined participants' memory predictions for massed and spaced items, following the work of Zechmeister and Shaughnessy (1980) and others (e.g., Dunlosky and Nelson 1994; Kornell and Bjork 2008; Kornell et al. 2010). In [Experiment 1](#), we examined this issue when participants made ratings after each presentation of the repeated (massed or spaced) items. In particular, participants might become more aware of the benefits of both repetition and spacing when these predictions are made for each presentation of the item (as opposed to only after the second presentation, as done in the study by Zechmeister and Shaughnessy 1980). In addition, participants also engaged in multiple study-test cycles (with unique lists of words), in order to determine if participants learned from experience with previous lists and recall tests and adjusted JOLs to better capture the benefits of spacing, to examine any effect of knowledge updating (e.g., Hertzog et al. 2009). In [Experiments 2](#) and [3](#), we attempted to make participants aware of the benefits of spacing by allowing them to score their own recall performance in terms of massed and spaced items that were recalled. Our reasoning was that this might draw attention to the differences in performance for spaced and massed items. Finally, in [Experiment 4](#), we directly informed participants about the benefits of spacing in memory performance, to compare how direct cueing influenced JOLs compared to potentially more subtle cues of experience and feedback. The results from these experiments are then discussed in terms of the cues that participants use when making JOLs for massed and spaced items, and how metacognitive predictions can be sensitive or insensitive to critical features that lead to the benefits of spacing in everyday learning.

Experiment 1

Although spacing enhances later recall performance, participants' predictions do not always reflect this finding, suggesting that people are not aware of the benefits of spaced rehearsal (Dunlosky and Nelson 1994; Zechmeister and Shaughnessy 1980). In [Experiment 1](#) we

sought to replicate the main finding in which recall, but not predictions of recall, were influenced by spacing of items. Participants studied a list of items, some of which were presented twice, in either a massed or spaced fashion. Upon the second presentation of these items, participants made a JOL. Based on previous work, we expected recall to favor spaced items, but that JOLs would be less sensitive to spacing.

We also were interested in whether participants could learn to assign higher JOLs to spaced items, with experience with multiple lists (study and tests), and making JOLs after each presentation of the item. This was done to draw participants' attention to the spaced and massed presentation of items in the context of memory predictions, and to allow for a measure of the change in JOLs from first to second presentation of the various items. Furthermore, several experiments have suggested that repeated testing improves JOL accuracy (e.g., Begg et al. 1989; King et al. 1980; Koriat 1997; Leonesio and Nelson 1990; Lovelace 1984). Thus, as done by others who have examined how experience with encoding and retrieval can influence the accuracy of subsequent metacognitive judgments (e.g., Castel 2008; Koriat 1997; Koriat and Bjork 2006; Price et al. 2008; Rhodes and Castel 2008a, b), participants were given experience with study and test conditions, in order to determine if JOLs would change with test experience. This process, referred to as knowledge updating, can lead to improvements in metacognitive accuracy in certain situations (Dunlosky and Hertzog 2000). Thus, the present study expanded on previous work by soliciting JOLs after each presentation of each item, as opposed to just the second presentation of each item (see Zechmeister and Shaughnessy 1980), as well as having participants engage in three study-test cycles with a unique list of words in each cycle, to determine if JOLs become more sensitive to the spacing effect with task experience.

Method

Participants Twenty-eight participants were recruited from undergraduate courses and received course credit for participating.

Apparatus The experiment was run on a Dell computer with a standard 15 in. monitor and implemented using E-prime software (Schneider et al. 2001). The stimuli were presented in the center of the screen, in white lowercase letters on a black background and printed in 18-point Arial font.

Materials The stimuli consisted of three lists of 19 medium frequency words selected from Kucera and Francis (1967). The word lists consisted of words that were presented only once (single presentation), words that were repeated immediately (massed presentation), and words that were repeated after a lag of three intervening items (spaced practice). Two words at the beginning and two words at the end of the list were used as buffer words to account for primacy and recency effects. The 15 remaining words—five in each condition (single, massed, spaced)—were counterbalanced across conditions.

Procedure Participants were asked to learn and recall three lists, and provided JOLs after each presentation of every item in each list. In the learning phase of each list, participants were told they would be given a list of words to study, one at a time, and that some words would be repeated at various points in the list. After the presentation of each word, they would be asked to predict how likely they would be to remember the word (also referred to as a judgment of learning, or JOL).

At the beginning of a trial in the learning phase, a word was presented on the computer screen for 6 s. After the second presentation of an item, once the word was cleared from the screen, participants were asked “How likely are you to remember this word?” and shown a rating scale from 0 % to 100 %, marked in 10 % increments. They entered their JOL from 0 % to 100 % using keys marked in the 10 % increments on the keyboard (0 %, 10 %, 20 %, etc.). The JOL question and scale remained on the screen for a total of 6 s or until a response was entered. After the JOL, there was a 500 ms fixation display in which a white crosshair appeared in the center of the screen before the next word was presented. After the learning phase, participants completed a distracter task that involved counting aloud backwards by 3’s for 30 s from a three-digit number. After the counting task was finished, participants were given instructions about the recall test. They were instructed to recall as many of the words from the list as possible by speaking them aloud to the experimenter, who recorded the participants’ responses on paper. After the learning and final recall phases for the first list of words, participants then proceeded to the learning and recall phases of the second list, followed by the learning and recall phases of the third list.

Results and discussion

The results from [Experiment 1](#) are displayed in [Fig. 1](#) (in terms of mean recall and JOLs for massed and spaced items as a function of list) and [Fig. 2](#) (mean recall and JOLs for the first and second presentation of massed and spaced items as a function of list).

Free recall As expected, there were significant effects of both repetition and spacing on free recall performance. Participants recalled more massed ($M=51\%$; $SE=4.1\%$) than once-presented items ($M=40\%$; $SE=4.2\%$), $F(1, 27)=9.89$, $MSE=466.31$, $p<.01$, $\eta_p^2=.27$ and recalled more spaced items ($M=67\%$; $SE=3.7\%$) than massed items, $F(1, 27)=37.51$,

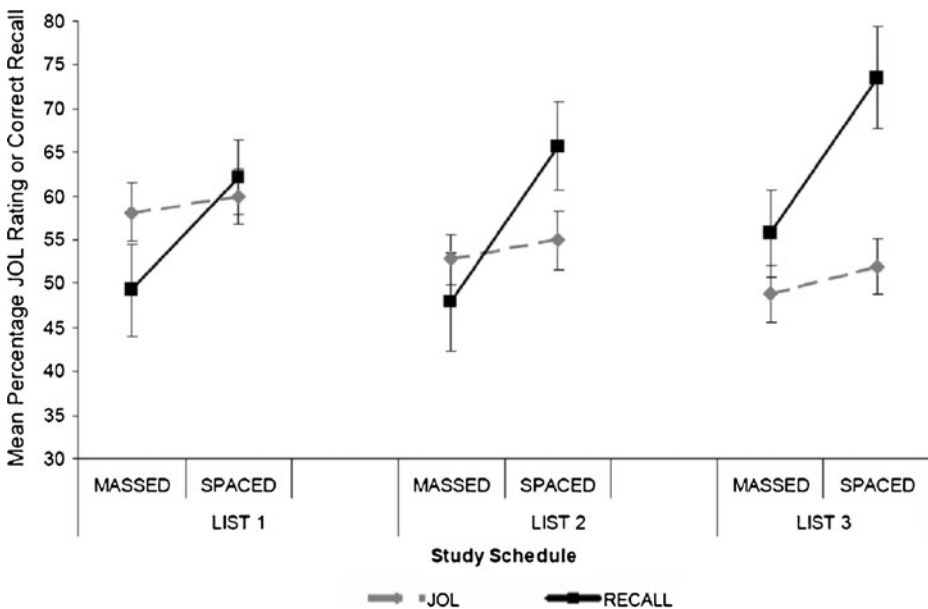


Fig. 1 Mean recall performance and judgment of learning (JOL) of spaced and massed items for each list in [Experiment 1](#). Error bars reflect standard error of the means in all figures

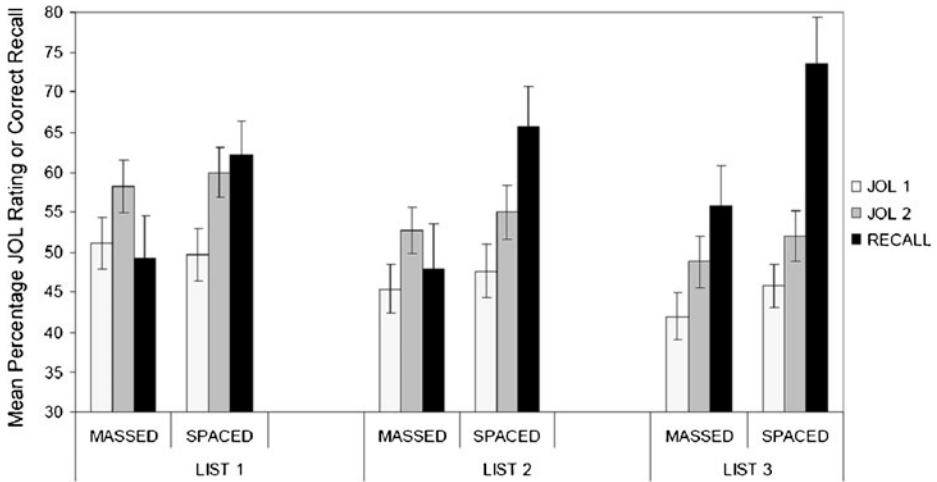


Fig. 2 Mean judgment of learning (JOL) for the first (JOL1) and second (JOL2) presentation of massed and spaced items, and mean recall for massed and spaced items, for each list in *Experiment 1*

$MSE=293.47$, $p<.001$, $\eta^2_p=.58$. There was also a slight trend towards a main effect of List, such that recall improved over lists, from 56 % in List 1 to 65 % in List 3, although this did not reach conventional levels of significance, $F(2, 54)=2.51$, $MSE=530.95$, $p=.09$, $\eta^2_p=.09$. This increasing trend in free recall across trials could be attributed to strategy changes across free-recall study/test cycles (e.g., Delaney and Knowles 2005). The List \times Spacing effect was not significant, $F<1$, as the mean spacing effect remained fairly stable across lists.

Judgments of learning JOLs were collected on all trials for all conditions, yielding one JOL for once presented items and two JOLs for items in the massed and spaced conditions. There was no reliable difference between JOLs for once-presented items and massed items on the first presentation ($F<1$). However, there was a significant difference between JOLs for once-presented items and the second presentation of massed items, such that participants gave higher JOLs to massed items on their second presentation ($M=53$ %; $SE=2.9$ %) than to the once-presented items ($M=46$ %; $SE=2.8$ %), $F(1, 27)=20.82$, $MSE=98.52$, $p<.001$, $\eta^2_p=.44$. Thus, participants increased their JOL ratings when an item was repeated. Differences in JOLs were also apparent based on study schedule (massed vs. spaced). Specifically, participants gave higher JOLs on the second presentation of an item ($M=54$ %; $SE=2.9$ %) compared to the first presentation of an item ($M=46$ %; $SE=2.8$ %), $F(1, 27)=27.51$, $MSE=171.78$, $p<.001$, $\eta^2_p=.51$. JOLs were also slightly but reliably higher for spaced items ($M=52$ %; $SE=2.8$ %) than massed items ($M=50$ %; $SE=2.8$ %), $F(1, 27)=8.88$, $MSE=36.08$, $p<.01$, $\eta^2_p=.25$, suggesting some awareness of the spacing effect. However, despite this slight increase in JOLs for spaced items, participants still underestimated the benefits of spacing by a substantial margin.

Comparison of recall with JOLs We carried out a direct comparison of JOLs and recall performance, in order to determine how these variables may or may not be related, and how this could change with task experience and/or knowledge updating. However, we note that these analyses should be treated with some caution as participants may use a restricted range when assigning JOLs, whereas recall performance is not limited by these scaling issues. In

terms of determining if Measure (JOL and recall) interacted with List (1st, 2nd, or 3rd) or Study Schedule (spaced or massed), a 2 (Measure: JOL, recall) \times 3 (List: first, second, third) \times 2 (Study Schedule: spaced, massed) repeated measures ANOVA was conducted for each JOL trial (initial judgment vs repeated judgment). For the initial JOL, there was a main effect of Measure such that actual recall was higher ($M=59\%$; $SE=3.7\%$) than JOLs ($M=47\%$; $SE=4.8\%$), $F(1, 27)=9.70$, $MSE=1270.31$, $p<.01$. $\eta^2_p=.26$. This difference between actual recall and JOLs increased across lists, as shown by a Measure \times List interaction. In particular, while recall increased over lists (55 % to 57 % to 65 % for Lists 1, 2 and 3, respectively), JOLs decreased across lists (50 % to 47 % to 44 %), $F(2, 54)=6.08$, $MSE=286.20$, $p<.01$, $\eta^2_p=.18$. This general decline of JOLs across multiple lists has been demonstrated in previous studies, and has been referred to as the “underconfidence with practice effect” (e.g., Koriat et al. 2002). As illustrated in Fig. 1, there was also a significant interaction between memory measure and study schedule, such that the spacing effect was much larger for recall (67 % vs 51 % for spaced and massed, respectively) than for initial JOLs (48 % vs 46 % respectively), $F(1, 54)=34.18$, $MSE=131.95$, $p<.001$, $\eta^2_p=.56$.

For the JOLs made on the second presentation of each item, there was a significant interaction between Measure and List, such that recall increased over list (55 % to 57 % to 65 %) but JOLs decreased over list (59 % to 54 % to 50 %), $F(2, 54)=7.69$, $MSE=290.84$, $p<.001$, $\eta^2_p=.22$. There was also a significant interaction between Measure and Study Schedule, such that the spacing effect was much larger in recall (67 % vs 51 % for spaced and massed, respectively) than in the repeated JOLs (57 % vs 53 % respectively), $F(1, 54)=25.89$, $MSE=154.94$, $p<.001$, $\eta^2_p=.49$. There was a much smaller disparity between JOLs for massed items ($M=53\%$; $SE=2.9\%$) and actual recall of massed items ($M=51\%$; $SE=4.1\%$), compared to JOLs given for spaced items ($M=56\%$; $SE=3.0\%$) and actual recall for the spaced items ($M=67\%$; $SE=3.7\%$). This suggests that although participants were quite accurate at predicting recall for massed items, they underestimated how likely they would be to recall the spaced items.

In an exploratory analysis designed to assess the potential influence of experience on participants’ JOL ratings over lists, we examined the correlation between the size of the spacing effect and difference in JOLs for massed vs spaced items for each participant. It may be the case that when participants show a sizable spacing effect, then JOLs reflect a difference between massed and spaced items. Alternatively, it could be the case that only those participants who show a spacing effect are aware of this difference, due to some participant characteristics. For List 3, the correlation between size of the spacing effect and size of the difference in JOLs for massed and spaced items was $r=-.02$, $p>.90$, indicating no relationship between the size of the spacing effect in actual memory performance and subjects’ JOLs for spaced vs massed items.

The results from Experiment 1 suggest that, although participants are well aware of the benefits of repetition and massed practice, JOLs were not highly sensitive to the effect of spaced presentation on later recall. Experiment 1 showed that participants’ JOLs did not differ for spaced compared to massed practice, whereas recall was better for spaced items. This replicates and extends the main findings from Zechmeister and Shaughnessy (1980) in which participants’ memory predictions failed to differentiate between massed or spaced items, despite actual memory performance being greater for spaced relative to massed items. In general, it appears that JOLs increased by approximately 10 percentage points from the first to the second presentation (massed or spaced), possibly reflecting an “anchoring and adjusting” mechanism (e.g., Scheck et al. 2004) that is common for both spaced and massed items. What is also present in the results is that overall recall increased with task experience (see also Delaney and Knowles 2005) while JOLs declined, which may reflect “underconfidence with practice” (Koriat et al. 2002). The main finding is that while JOLs were fairly well calibrated for massed items, JOLs underestimated actual recall for spaced items.

Experiment 2

One potential reason that participants' memory predictions may not accurately capture the role of spacing in recall may be the lack of feedback regarding recall performance for massed and spaced items. For example, participants may not understand the impact of spacing because, in a free recall task, they are unable to distinguish between those items that are recalled which were presented in a spaced compared to massed fashion. Thus, if participants were aware that they recalled a larger number of spaced items, they might adjust their JOLs accordingly. **Experiment 2** addressed this issue by providing specific feedback on recall output. The procedure was identical to **Experiment 1**, with the exception that participants were informed of their performance after each list. Specifically, in order to make participants aware of their own recall performance, participants in **Experiment 2** scored their own recall immediately after the recall session (see also Rawson and Dunlosky 2007), and tabulated the number of spaced, massed, and single items that were recalled. They then engaged in a second and third list of unique items, and continued to score their own recall output after each recall trial. Under these conditions, of specific interest was whether participants' JOLs would reflect their performance for spaced and massed items given previous experience and awareness of differences in recall performance for spaced, massed and single items.

Method

Participants Forty-seven participants were recruited from undergraduate courses and received course credit for participating.

Apparatus, materials, and procedure These were identical to **Experiment 1**, with one exception: participants were given feedback on their performance after every list by scoring their own recall sheets. After the recall phase of each list, participants were given a sheet that listed the words they had just been asked to learn, divided according to condition, with the labels "spaced practice," "massed practice," "studied once," "beginning of list," and "end of list." Participants were instructed as to what each label meant. Using their recall sheet, participants were instructed to give themselves one point for every word they correctly recalled. They then wrote down the number of points they received in each condition. After grading their recall sheets, they then proceeded to the next list of words until they had studied and been tested on a total of three lists.

Results and discussion

The results from **Experiment 2** are displayed in Fig. 3 (in terms of overall recall and JOLs for massed and spaced items as a function of list) and Fig. 4 (recall and JOLs for the first and second presentation of massed and spaced items as a function of each list).

Free recall As expected, there were significant effects of both repetition and spacing on free recall performance. Participants remembered massed items ($M=53\%$; $SE=2.9\%$) better than once-presented items ($M=45\%$; $SE=3.0\%$), $F(1, 46)=12.09$, $MSE=394.82$, $p<.01$, $\eta^2_p=.21$, and recalled more spaced items ($M=67\%$; $SE=3.1\%$) than massed items, $F(1, 46)=33.21$, $MSE=444.40$, $p<.001$, $\eta^2_p=.42$.

Judgments of learning As in **Experiment 1**, JOLs were collected on all trials for all conditions, yielding one JOL for once presented items and two JOLs for items in the massed

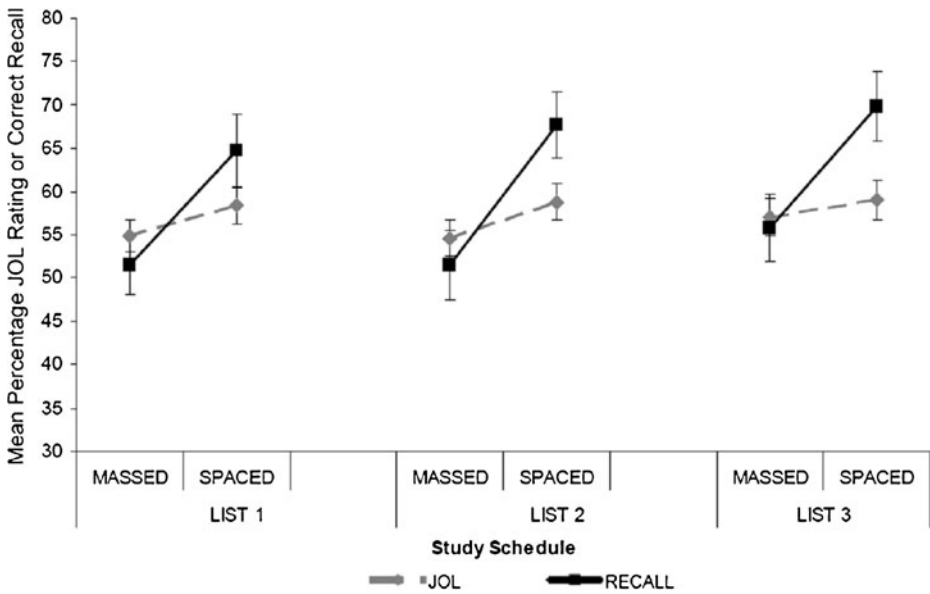


Fig. 3 Mean recall performance and judgment of learning (JOL) of spaced and massed items for each list in Experiment 2

and spaced conditions. In terms of repetition effects in JOLs, there was no reliable difference between JOLs for once-presented items and the initial presentation of the massed or spaced items ($F < 1$). However, there was a significant difference between JOLs for once-presented items and the repeated presentation of massed items, such that participants gave higher JOLs to massed items on their second presentation ($M = 56\%$; $SE = 1.8\%$) than to the once-presented items ($M = 47\%$; $SE = 1.7\%$), $F(1, 46) = 78.82$, $MSE = 70.65$, $p < .001$, $\eta^2_p = .63$.

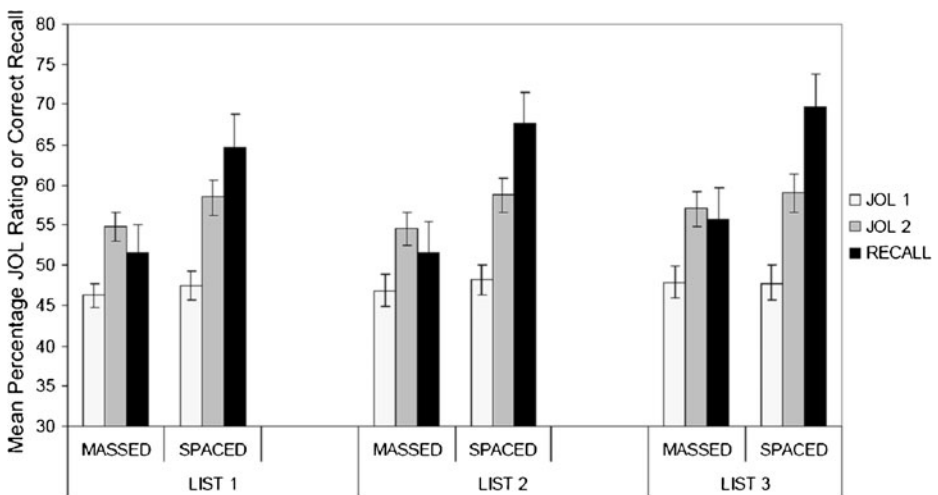


Fig. 4 Mean judgment of learning (JOL) for the first (JOL1) and second (JOL2) presentation of massed and spaced items, and mean recall for massed and spaced items, for each list in Experiment 2

Thus, participants increased their JOL ratings when an item was repeated. In general, participants gave higher JOLs on the second presentation of an item ($M=57\%$; $SE=1.8$) compared to the first presentation of an item ($M=47\%$; $SE=1.6\%$), $F(1, 46)=118.64$, $MSE=112.10$, $p<.001$, $\eta^2_p=.72$. There was no difference in JOLs for spaced and massed items on the initial JOL (47% vs 48%, respectively) but there was a difference on the repeated JOL (59% vs 56%, for spaced vs. massed, respectively) Thus, participants were not only reliably giving higher JOLs to repeated items, they were correctly judging spaced items as more likely to be remembered compared to massed items, although they still greatly underestimated actual performance for spaced items.

Comparison of recall with JOLs The relation between recall and JOLs was examined in a 2 (Measure) x 3 (List) x 2 (Study Schedule) repeated measures ANOVA each JOL trial (initial judgment vs repeated judgment). For the initial JOL, there was a main effect of Measure such that actual recall was higher ($M=60\%$; $SE=2.7\%$) than JOLs ($M=47\%$; $SE=1.6\%$), $F(1, 46)=21.42$, $MSE=1068.46$, $p<.001$, $\eta^2_p=.32$. As illustrated in Fig. 3, there was also a significant interaction between Measure and Study Schedule, $F(1, 46)=27.17$, $MSE=241.54$, $p<.001$, $\eta^2_p=.37$, such that the spacing effect was much larger in recall (67% vs 53% for spaced and massed, respectively) than in initial JOLs (48% vs 47% for spaced and massed, respectively).

For the repeated JOL, there was also a significant interaction between Measure and Study Schedule, $F(1, 46)=19.85$, $MSE=224.74$, $p<.001$, $\eta^2_p=.30$, such that the spacing effect was much larger in recall (67% vs 53% for spaced and massed, respectively) than in the repeated JOLs (59% vs 56% respectively). As in Experiment 1, there was a much smaller disparity between JOLs for massed items ($M=56\%$; $SE=1.8\%$) and actual recall of massed items ($M=53\%$; $SE=2.9\%$), whereas JOLs given for spaced items ($M=59\%$; $SE=1.9\%$) were significantly lower than actual recall for the spaced items ($M=67\%$; $SE=3.1\%$). Thus, participants were quite accurate at predicting recall for massed items, and reliably increased their JOLs for spaced compared to massed items upon repetition, but they still underestimated how likely they would be to recall the spaced items within each list. Thus, the impact of spacing on recall was much more substantial than the impact of spacing on JOLs.

As in Experiment 1, the correlation between a subjects' spacing effect in actual recall and the spacing effect in their JOLs was computed for the final list, List 3. Contrary to Experiment 1, there was a significant correlation between the size of the spacing effect in memory and the size of the spacing effect in JOLs, $r(47)=.34$, $p<.02$. Individuals with larger spacing effects were more likely to adjust their JOLs for spaced and massed items accordingly.

Experiment 3

In order to determine if participants may become sensitive to spacing when a larger spacing effect is present, we conducted an experiment in which the lag between spaced items was increased, as well as the total number of items in the list. Increasing the lag between spaced items should lead to a greater spacing effect (e.g., Glenberg 1976, 1977; Madigan 1969; Melton 1970; Murdock 1974; Underwood 1969). This manipulation was used to allow for a greater number of spaced items to be recalled at test, which may then draw attention to the presence of a spacing effect when participants score their own recall output. Thus, a lag of eight items between spaced repetitions was used, as compared to a lag of four items in the previous experiments. The combination of a greater lag and longer lists enhanced the overall spacing effect.

Method

Participants Twenty-four participants were recruited from undergraduate courses and received course credit for participating.

Apparatus, materials, and procedure These were identical to [Experiment 2](#), with two key exceptions. A longer lag was used for the spacing condition in order to produce a larger spacing effect, and a longer list was used in an effort to make the spacing between items more salient to participants. A lag of 8 items between spaced repetitions was used, as compared to a lag of 4 items in the previous experiments. In addition, the number of words per condition was increased from five to eight, which increased list length from 29 words in [Experiments 1 and 2](#) to 45 words, so subjects had more experience with all conditions.

Results and discussion

The results from [Experiment 3](#) are displayed in [Fig. 5](#) (in terms of overall recall and JOLs for massed and spaced items as a function of list) and [Fig. 6](#) (recall and JOLs for the first and second presentation of massed and spaced items as a function of each list).

Free recall As expected, there were significant effects of both repetition and spacing on free recall performance. Participants remembered massed items ($M=56\%$; $SE=3.1\%$) better than once-presented items ($M=56\%$; $SE=3.1\%$), $F(1, 23)=12.0$, $MSE=314.72$, $p<.01$, $\eta^2p=.34$, and recalled more spaced items ($M=70\%$; $SE=3.5\%$) than massed items, $F(1, 23)=35.47$, $MSE=226.26$, $p<.001$, $\eta^2p=.60$.

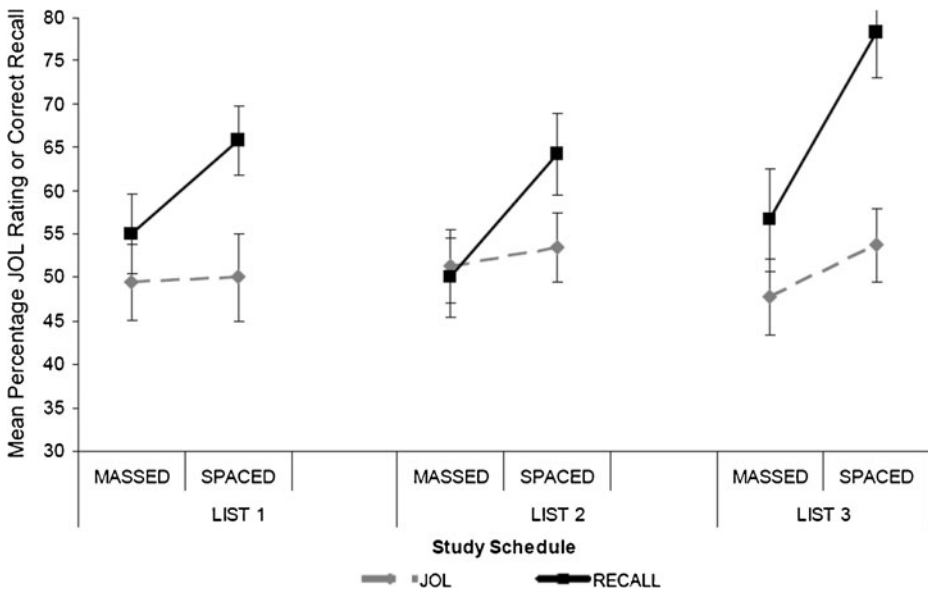


Fig. 5 Mean recall performance and judgment of learning (JOL) of spaced and massed items for each list in [Experiment 3](#)

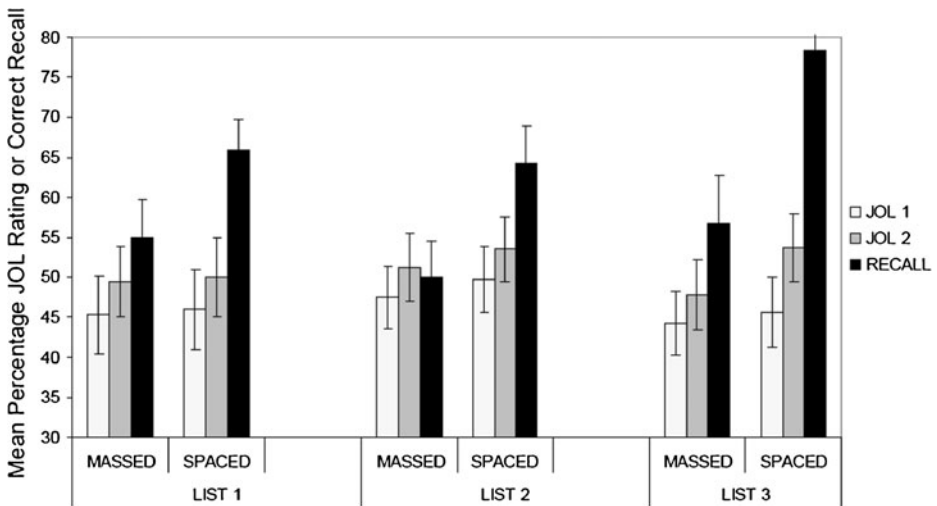


Fig. 6 Mean judgment of learning (JOL) for the first (JOL1) and second (JOL2) presentation of massed and spaced items, and mean recall for massed and spaced items, for each list in [Experiment 3](#)

Judgments of learning As in [Experiment 2](#), JOLs were collected on all trials for all conditions, yielding one JOL for once presented items and two JOLs for items in the massed and spaced conditions. In terms of repetition effects in JOLs, there was no reliable difference between JOLs for once-presented items and the initial presentation of the massed or spaced items ($F < 1$). There was a significant difference when comparing JOLs for once-presented items and the repeated presentation of massed items, however, such that JOLs were higher for massed items upon repetition (56 %) compared to once-presented items (51 %), $F(1, 23) = 8.50$, $MSE = 84.32$, $p < .01$, $\eta^2 p = .27$, indicating that JOLs were sensitive to repetition. There was no difference in JOLs for spaced and massed items on the initial JOL (51.5 % vs 50.6 %, respectively). For the repeated JOL, however, JOLs were significantly higher for spaced (58.4 %) compared to massed (55.7 %) items, $F(1, 23) = 5.10$, $MSE = 48.47$, $p < .04$, $\eta^2 p = .18$, indicating that participants were somewhat sensitive to the benefits of spacing in recall.

Comparison of recall with JOLs The relation between recall and JOLs was examined in a 2 (Measure) \times 3 (List) \times 2 (Study Schedule) repeated measures ANOVA for each JOL trial (initial judgment vs repeated judgment). For the initial JOL, as shown in [Fig. 6](#), there was a significant interaction between Measure and Study Schedule, $F(1, 23) = 25.89$, $MSE = 136.46$, $p < .001$, $\eta^2 p = .53$, such that the spacing effect was much larger in recall (70 % vs 56 % for spaced and massed, respectively) than in initial JOLs (51.5 % vs 50.6 % for spaced and massed, respectively).

For the repeated JOL, there was also a significant interaction between Measure and Study Schedule, $F(1, 23) = 23.29$, $MSE = 117.11$, $p < .001$, $\eta^2 p = .50$, such that the spacing effect was much larger in recall (70.5 % vs 55.6 % for spaced and massed, respectively) than in the repeated JOLs (58.4 % vs 55.7 % respectively). As in [Experiment 2](#), there was a much smaller disparity between JOLs for massed items ($M = 55.6$ %; $SE = 3.1$ %) and actual recall of massed items ($M = 55.7$ %; $SE = 2.0$ %), whereas JOLs given for spaced items ($M = 58.4$ %; $SE = 2.1$ %) were significantly lower than actual recall for the spaced items ($M = 70.4$ %; $SE =$

3.5 %). Thus, participants were quite accurate at predicting recall for massed items, and reliably increased their JOLs for spaced compared to massed items upon repetition, but they still underestimated how likely they would be to recall the spaced items within each list. As in previous experiments, the impact of spacing on recall was much more substantial than the impact of spacing on JOLs.

As in the previous experiments, the correlation between subjects' spacing effect in actual recall and the spacing effect in JOLs was computed for the final list, List 3. There was a marginally significant correlation between the size of the spacing effect in memory and the size of the spacing effect in JOLs, $r(24) = .39, p = .06$, similar to [Experiment 2](#).

Results from [Experiment 3](#) were similar to those in previous experiments. JOLs for massed items were very close to actual recall for massed items, but JOLs for spaced items still significantly underestimated the benefits of spacing to final recall. This was the case despite the fact that lists were changed to make the manipulation of spacing more salient to participants, by increasing the number of spaced and massed items and doubling the lag between spaced items from 4 to 8. Again, as in [Experiment 2](#), by the final list, there was a correlation between the magnitude of participants' spacing effects in recall and the spacing effect reflected in their JOLs. This may indicate that participants' awareness of the benefits of spacing on memory are most likely to occur when the benefits of spacing are particular obvious and apparent. In these experiments, there is an indication that this awareness is developing after experience with multiple lists, but it is still not enough to greatly increase subsequent JOLs for spaced versus massed items by the final list.

Experiment 4

The findings thus far indicate that, although participants are quite accurate at predicting their performance for repeated items in a massed condition, they are still not sensitive enough to the spacing effect to adjust their JOLs for spaced and massed items appropriately, despite increased experience and feedback with the spacing effect. More recent work has shown that some knowledge updating can occur with experience, but only when trials are blocked, making them more apparent to the learner. Price et al. (2008) suggest that one reason JOLs may not show substantial updating after task experience is that participants do not link performance at test with previous encoding operations. To make the link more obvious, Price et al. blocked strategy type at test (e.g., imagery items were tested in a separate block from repetition items). In this case, JOLs showed somewhat better sensitivity to strategy type. Of course, blocking trials is impossible in the case of spacing, by definition, so other ways to draw attention to the benefits of spacing are needed in multi-trial designs. The previous experiments examined this by allowing for self-scoring of recall performance, which led to a trend in increased JOLs for spaced items in subjects' showing large spacing effects, but perhaps instituting some factor at encoding that makes spacing salient would allow for greater learning about the effect.

[Experiment 4](#) was conducted to explore the role that more direct awareness of the spacing effect in recall could have on subsequent JOLs. Participants first studied a word list and gave JOLs for each item and received feedback on their performance for each class of items after recall, as in previous experiments. Unlike the previous studies, however, after the first list, participants were given instruction on the spacing effect, including what it was and how spacing could benefit their memory. Then, they studied another list of words similar to the first list, this time armed with direct knowledge of the spacing effect. The effect of this direct awareness on JOLs and subsequent recall was the focus of this experiment.

Method

Participants Twenty-five participants were recruited from undergraduate courses and received course credit for participating.

Apparatus, materials, and procedure These were similar to [Experiment 3](#), with two key exceptions. Only two lists of to-be-remembered items were used in [Experiment 4](#). The lists were the same as those used in [Experiment 3](#), with a lag of eight items, with eight items per condition, etc. The procedure for each list was also the same as [Experiment 3](#): Participants studied and made JOLs for each word as it appeared, recalled the words, and graded their recall sheet afterwards. However, before beginning List 2, participants were given information on the spacing effect. After being told that they would now study and rate a new list of words, they read the following instructions on the screen just before List 2 was presented:

“You may have noticed that some of the words are repeated - sometimes you see a word twice in a row (MASSED STUDY), other times it is repeated a little later in the list (SPACED STUDY). Researchers know that repetition is good for memory, and spaced repetition tends to be much better than massed repetition. The advantage for spaced items compared to massed items is called the Spacing Effect. You may see a Spacing Effect in your own memory if your SPACED STUDY score is higher than your MASSED STUDY score on your grading sheet. The Spacing Effect can be a very powerful effect in memory, but people tend to underestimate how much spacing will help their memory. Keep the Spacing Effect in mind as you study and rate the next words for your memory test!”

After the second list was presented for study and JOLs, participants again recalled the items and graded their recall sheets.

Results & discussion

The results from [Experiment 4](#) are displayed in [Fig. 7](#) (in terms of overall recall and JOLs for massed and spaced items as a function of list) and [Fig. 8](#) (recall and JOLs for the first and second presentation of massed and spaced items as a function of each list).

Free recall As expected, there were significant effects of both repetition and spacing on free recall performance. Participants remembered massed items ($M=48\%$; $SE=3.7\%$) better than once-presented items ($M=40\%$; $SE=3.6\%$), $F(1, 24)=6.32$, $MSE=269.27$, $p<.02$, $\eta^2p=.21$, and recalled more spaced items ($M=69\%$; $SE=2.9\%$) than massed items, $F(1, 24)=37.70$, $MSE=285.55$, $p<.001$, $\eta^2p=.61$. There was also a significant List x Spacing interaction, $F(1, 24)=5.11$, $MSE=293.88$, $p<.04$, $\eta^2p=.18$, such that the spacing effect was larger on List 2 (29 %) than List 1 (13 %).

Judgments of learning As in [Experiments 2 and 3](#), JOLs were collected on all trials for all conditions, yielding one JOL for once presented items and two JOLs for items in the massed and spaced conditions. In terms of repetition effects in JOLs, there was a small but reliable difference between JOLs for once-presented items and the initial presentation of the massed or spaced items, $F(1, 24)=8.13$, $MSE=24.53$, $p<.01$, $\eta^2p=.25$, such that initial JOLs for massed items (45 %) were actually slightly smaller than JOLs for once-presented items (47 %). However, when comparing JOLs for once-presented items and the repeated

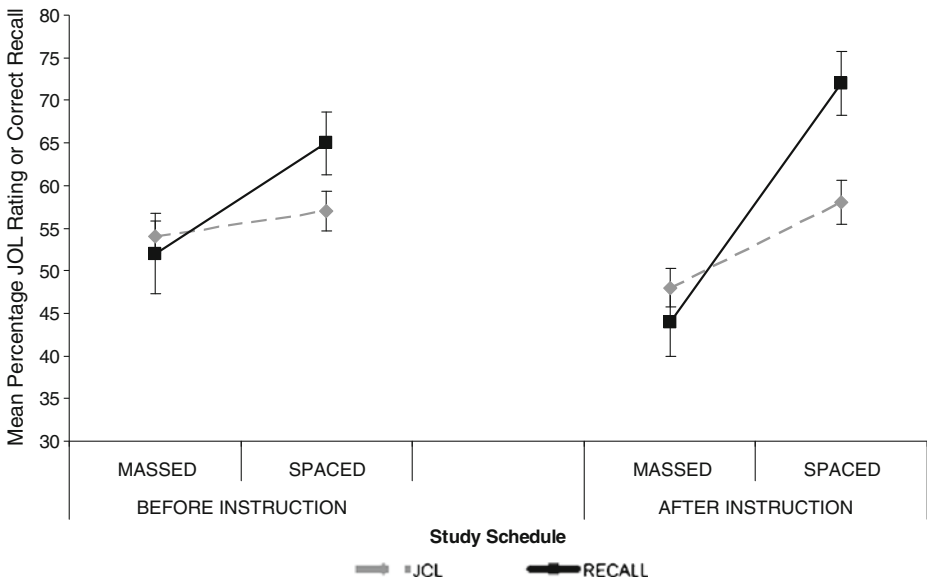


Fig. 7 Mean recall performance and judgment of learning (JOL) of spaced and massed items before and after instructions on the spacing effect in Experiment 4

presentation of massed items, JOLs for massed items (51 %) were significantly higher than once-presented items, $F(1, 24)=11.16$, $MSE=21.11$, $p<.01$, $\eta^2p=.32$, indicating that JOLs were sensitive to repetition. There was no difference in JOLs for spaced and massed items on the initial JOL (48 % vs 46 %, respectively). However, for the repeated JOL, there was a significant difference between spaced (58 %) and massed items (51 %), $F(1, 24)=33.75$,

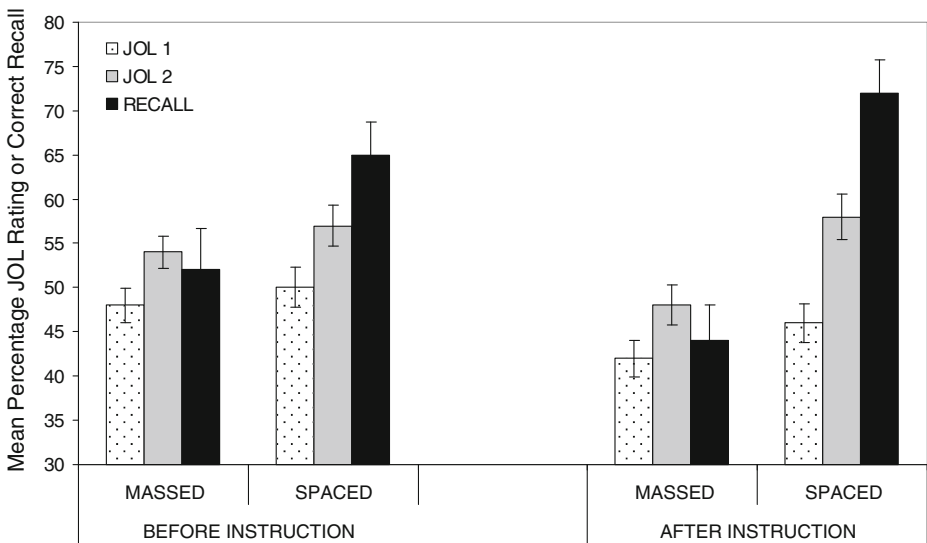


Fig. 8 Mean judgment of learning (JOL) for the first (JOL1) and second (JOL2) presentation of massed and spaced items, and mean recall for massed and spaced items, before and after instructions on the spacing effect in Experiment 4

$MSE=33.51$, $p<.001$, $\eta^2p=.58$. There was also a significant List x Study Schedule interaction, $F(1, 24)=7.58$, $MSE=33.54$, $p<.02$, $\eta^2p=.24$, which showed that the spacing effect in JOLs was greater in List 2 (10 %), after instructions on the spacing effect, than List 1 (3 %). Thus, direct information on the spacing effect appeared to increase awareness of the benefits of spacing and subsequently influenced JOLs for spaced versus massed items.

Comparison of recall with JOLs The relation between recall and JOLs was examined in a 2 (Measure) x 3 (List) x 2 (Study Schedule) repeated measures ANOVA for each JOL trial (initial judgment vs repeated judgment). For the initial JOL, as shown in Fig. 8, there was a significant interaction between Measure and Study Schedule, $F(1, 24)=24.81$, $MSE=155.73$, $p<.001$, $\eta^2p=.51$, such that the spacing effect was much larger in recall (69 % vs 48 % for spaced and massed, respectively) than in initial JOLs (48 % vs 46 % for spaced and massed, respectively).

For the repeated JOL, there was also a significant interaction between Measure and Study Schedule, $F(1, 24)=16.04$, $MSE=153.28$, $p<.01$, $\eta^2p=.40$, such that the spacing effect was still larger in recall (69 % vs 48 % for spaced and massed, respectively) than in the repeated JOLs (58 % vs 51 % respectively). As in Experiment 3, there was a much smaller disparity between JOLs for massed items ($M=51$ %; $SE=1.7$ %) and actual recall of massed items ($M=48$ %; $SE=3.7$ %), whereas JOLs given for spaced items ($M=58$ %; $SE=2.3$ %) were significantly lower than actual recall for the spaced items ($M=69$ %; $SE=2.9$ %). Thus, participants were more accurate at predicting recall for massed items than spaced items, although they did reliably increase their JOLs for spaced compared to massed items upon repetition after instructions on the spacing effect. Thus, instruction on the spacing effect served to influence JOLs for spaced and massed items to reflect a greater awareness of the spacing effect, but JOLs for spaced items still underestimated actual recall of spaced items. This is especially striking given the instructional information provided prior to the study, and may suggest that participants use more cue-based information, relative to theory-based information, when making JOLs regarding massed and spaced items. Given that this effect did not change substantially across lists, a lack of knowledge updating was observed that is consistent with prior studies (e.g., Hertzog et al. 2009; Price et al. 2008), and suggests that learners may be more focused on specific properties of the information, and thus do not incorporate the benefits of temporal spacing when making item-based JOLs.

General discussion

The findings from the present studies show that although JOLs are highly sensitive to the repetition of items at encoding, JOLs do not accurately capture the benefits of spaced rehearsal (see also Dunlosky and Nelson 1994; Kornell and Bjork 2008; Zechmeister and Shaughnessy 1980). This effect persisted despite participants making JOLs during each presentation of the item, and even when participants were made aware of the presence of a spacing effect by scoring their own recall prior to engaging in a second similar memory task. Although JOLs did show a slight trend in terms of differentiating between spaced and massed practice, the JOLs were much more accurate for massed than spaced items, as participants greatly underestimated the recall of spaced items. Thus, although participants may not appreciate the benefits of spacing, this type of observation allows for a better understanding of cues that are used when making judgments about memory performance, and why certain cues (repetition) are given more weight than others (schedule of repetition) when making judgments of learning.

Our findings may be consistent with Koriat's (1997) cue-utilization approach, which argues that extrinsic cues, such as repetition, are discounted when making JOLs. Recent research has shown that participants may underestimate the benefits of repeated study on later learning (Kornell and Bjork 2009). The present findings suggest that participants may take repetition into account but show less sensitivity to the benefits of spaced repetition. One explanation for the current results is that participants use ease of processing or fluency when making JOLs (e.g., Begg et al. 1989; Castel et al. 2007; Yue et al. 2012), and repeated, massed items are perceived as more fluent than spaced items, despite later recall favoring spaced items. Specifically, in the present study, participants may report that if they just studied some information, an immediate presentation of this same information should enhance learning. However, fluency might have a slightly different effect when items are presented within a longer temporal context (i.e., spaced). In this type of situation, this may lead to a "discounted JOL", in which participants may rely on some sort of anchoring and adjustment mechanism (cf. Scheck et al. 2004). With massed items, the initial JOL at first presentation serves as an anchor, and with the second immediate presentation, participants then "logically" enhance their ratings. However, with spacing, upon the second spaced presentation of an item, the participant may then be reminded that this information was in fact presented earlier, but was possibly not well encoded (due somewhat less fluent processing of spaced items as compared to fluency that is accrued in massed fashion). Although somewhat speculative, this experience and insight regarding the current status of the item in memory may then cause the participant to adjust their JOL to capture this intuition, adjusting downward from what might actually have been an accurate initial anchor.

Several additional theoretical explanations may exist regarding why participants do not give accurate JOLs for spaced items. One possibility may be based on encoding variability, in that participants have stronger access to the present and prior instances of massed items, and weaker access to the first item instances for spaced items when making JOLs. Thus, accessibility plays a dominant role when making JOLs, whereas variability may be a better diagnostic cue, but is less accessible. Another possibility is that participants understand that massed presentations lead to good performance for immediate tests, and do not anticipate a long-term memory test, in which spaced items are better recalled relative to massed items. However, if this were the case, then with task experience (e.g., on List 2 or 3), participants should be aware of the retrieval conditions, and it appears that participants do not show knowledge updating (see also Hertzog et al. 2009; Price et al. 2008). While the present results do not provide a firm delineation between various theoretical accounts, the observation of learning about the benefits of spacing with task experience are important to better understand how experience-based learning can inform item-based metacognitive judgments.

It should be noted that the present experiment represents an approach to the study of spacing that may be somewhat far removed from actual study and testing that occurs in the classroom, in which longer and more variable retention intervals exist and richer materials are studied and retrieved. Although the benefits of spacing are widely known to cognitive psychologists, educators, and possibly many students, in the present experiments, students may employ a different "mental model" of how information is retained and forgotten. While the hallmark of learning is often evidenced by tests of long-term memory, students may feel that learning or mastery can be quickly assessed by an immediate test, and in fact, in terms of short-term retention (Bjork 1999), massed practice can lead to better memory performance than spacing. The present study attempted to give participants a certain amount of experience with study-test episodes that lead to a robust spacing effect, but perhaps JOLs are more a measure of immediate processing fluency (Begg et al. 1989) and not a reflection of how encoding variability and retention interval might influence learning.

In general, students may be most comfortable with massed practice, and for this reason their predictions regarding the effect of repetition and massed practice are in fact quite accurate. However, spacing presents a situation in which encountering previously studied material may lead to some discrepancies regarding actual learning and assessment of the desired level of learning (cf. Dunlosky and Hertzog 1998; Thiede and Dunlosky 1999). Recent work by Price et al. (2008) suggests that one reason JOLs may not show substantial updating after task experience is that participants do not link performance at test with previous encoding operations (see also Kornell and Bjork 2009). Price et al. (2008) blocked strategy type at test so that imagery items were tested in a separate block from repetition items. Blocking strategy type led to somewhat better sensitivity to strategy type in JOLs. In the present study, we attempted to draw attention to the benefits of spacing by using a multi-trial design that allowed for experience and feedback about spacing effects through self-scoring of recall performance, which led to small but reliable trends in increased JOLs for spaced items in subjects showing large spacing effects. In our final experiment, [Experiment 4](#), we used perhaps the most obvious method possible to draw attention to the benefits of spacing, by directly instructing students about these benefits and how spacing could influence their subsequent memory. This direct instruction led to the largest effects in JOLs, a full 10 % difference between JOLs for spaced and massed items after instruction. Although JOLs in this experiment still underestimated the magnitude of the spacing effect in recall, they were nonetheless influenced more substantially by instruction about the spacing effect than by the variables of repetition and feedback.

Although participants did not appear to adequately assess how spacing would influence recall, they were quite attuned to how repetition in general would influence recall—a factor that is somewhat more accessible than presentation schedule. However, it may be difficult to differentiate how repetition and spacing can influence recall using common metacognitive index, such as a JOL. It may be the case that when making JOLs, participants place a priority on the effect of one variable (repetition: presented once or twice), and can adequately incorporate this into JOLs, with other ostensibly “secondary variables” either ignored or treated as noise. This would allow for the development and use of a simple heuristic based on repetition. Interestingly, there is evidence that participants attempted to refine the use of this heuristic as they became more aware of the benefits of spacing, by recalling more spaced items than massed items, especially when given instruction, although this effect was limited when compared to the actual performance effect of spacing. While students can often appreciate the benefits of spacing in theory, and report using this schedule when studying for tests and exams, it is important to note that, based on the present results, students may indeed underestimate the potential of spacing in a variety of settings if they are not keenly aware of its benefits.

The present findings have implications for how students study material, especially in self-regulated learning environments. In the context of monitoring learning, participants in the present studies began to appreciate the benefits of spacing, but still under-estimated these benefits despite direct instruction regarding how spacing can help memory. It may be that students typically use a “shorter-term” memory model when assessing the potential benefits of massing and spacing and under short-term memory conditions, massing or cramming can lead to good memory performance, but not necessarily longer-term learning (see also Koriari et al. 2004; Kornell and Bjork 2009). Students may begin to appreciate the benefits of spacing when material differs in terms of the difficulty in initial learning (e.g., Benjamin and Bird 2006) and the test is based on long-term learning. Our findings showed that a combination of instruction, experience, and feedback produced the largest spacing effects in JOLs, revealing that students are indeed sensitive to this information. Outside of the

laboratory, students may learn to appreciate the benefits of spacing more quickly and adapt their own study strategies accordingly when such instruction, experience, and feedback are given to them in the context of impacting their grades (Toppino and Cohen 2010). Given that spacing can have such a powerful effect on retention, it would also be of interest to extend these results using more educationally-relevant materials (or in actual classrooms) and using a more diverse range of learners rather than the typical college student (Logan and Balota 2008; Meyer and Logan, *in press*). Overall, since students can engage in a variety of learning techniques when they self-pace their own study, future research is needed to develop ways for students to discover and understand how spacing can enhance both memory performance and long-term learning in self-regulated learning environments.

References

- Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: when agendas override item-based monitoring. *Journal of Experimental Psychology: General*, *138*, 432–447.
- Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, *21*, 627–635.
- Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*, 610–632.
- Benjamin, A. S., & Bird, R. D. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language*, *55*, 126–137.
- Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). Cambridge: MIT Press.
- Castel, A. D. (2008). Metacognition and learning about primacy and recency effects in free recall: the utilization of intrinsic and extrinsic cues when making judgments of learning. *Memory & Cognition*, *36*, 429–437.
- Castel, A. D., McCabe, D. P., & Roediger, H. L., III. (2007). Illusions of competency and overestimation of associative memory for identical items: evidence from judgments of learning. *Psychonomic Bulletin & Review*, *14*, 107–111.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–380.
- Delaney, P. F., & Knowles, M. E. (2005). Encoding strategy changes and spacing effects in the free recall of unmixed lists. *Journal of Memory and Language*, *52*, 120–130.
- deWinstanley, P. A., & Bjork, E. L. (2004). Processing strategies and the generation effect: implications for making a better reader. *Memory & Cognition*, *32*, 945–955.
- Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah: Erlbaum.
- Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: a componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging*, *15*, 462–474.
- Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language*, *33*, 545–565.
- Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, *15*, 1–16.
- Glenberg, A. M. (1977). Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 282–294.
- Hertzog, C., Price, J., Burpee, A., Frentzel, W. J., Feldstein, S., & Dunlosky, J. (2009). Why do people show minimal knowledge updating with task experience: inferential deficit or experimental artifact? *Quarterly Journal of Experimental Psychology*, *62*, 155–173.
- Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of comprehensions and competence. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing individual and team performance* (pp. 57–80). Washington, DC: National Academy Press.
- King, J. F., Zechmeister, E. B., & Shaughnessy, J. J. (1980). Judgments of knowing: the influence of retrieval practice. *The American Journal of Psychology*, *93*, 329–343.

- Koriat, A. (1997). Monitoring one's own knowledge during study: a cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349–370.
- Koriat, A., & Bjork, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition*, *34*, 959–972.
- 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*, 147–162.
- Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. (2004). Predicting one's own forgetting: the role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, *133*, 643–656.
- Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: is spacing the “enemy of induction”? *Psychological Science*, *19*, 585–592.
- 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.
- Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, *25*, 498–503.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 464–470.
- Logan, J. M., & Balota, D. A. (2008). Expanded vs. equal interval spaced retrieval practice: exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, & Cognition*, *15*, 257–280.
- Lovelace, E. A. (1984). Metamemory: monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 756–766.
- Madigan, S. A. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, *8*, 828–835.
- Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 596–606.
- Meyer, A. N. D., & Logan, J. M. (in press). Taking the testing effect beyond the college freshman: benefits for lifelong learning. *Psychology and Aging*.
- Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.
- Price, J., Dunlosky, J., & Hertzog, C. (2008). Age-related differences in strategy knowledge updating: blocked testing produces greater improvements in metacognitive accuracy for younger than older adults. *Aging, Neuropsychology, and Cognition*, *15*, 601–626.
- Pyc, M. A., & Dunlosky, J. (2010). Toward an understanding of students' allocation of study time: Why do they decide to mass or space their practice? *Memory and Cognition*, *38*, 431–440.
- Rawson, K., & Dunlosky, J. (2007). Improving students' self-evaluation of learning for key concepts in textbook materials. *European Journal of Cognitive Psychology*, *19*, 559–579.
- Rhodes, M. G., & Castel, A. D. (2008a). Metacognition and part-set cuing: can interference be predicted at retrieval? *Memory & Cognition*, *36*, 1429–1438.
- Rhodes, M. G., & Castel, A. D. (2008b). Memory predictions are influenced by perceptual information: evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*, 615–625.
- Scheck, P., Meeter, M., & Nelson, T. O. (2004). Anchoring effects in the absolute accuracy of immediate versus delayed judgments of learning. *Journal of Memory and Language*, *51*, 71–79.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2001). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools.
- Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 907–912.
- Son, L. K. (2004). Spacing one's study: evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 601–604.
- 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.
- Toppino, T. C., & Cohen, M. S. (2010). Metacognitive control and spaced practice: Clarifying what people do and why. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1480–1491.

- Toppino, T. C., Cohen, M. S., Davis, M. L., & Moors, A. C. (2009). Metacognitive control over the distribution of practice: when is spacing preferred? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1352–1358.
- Underwood, B. J. (1969). Attributes of memory. *Psychological Review*, *76*, 559–573.
- Yue, C. L., Castel, A. D., & Bjork, R. A. (2012). When disfluency is - and is not - a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition*.
- Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you think you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society*, *15*, 41–44.