Cognition 205 (2020) 104427

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Part-list cuing with prose material: When cuing is detrimental and when it is not

Lisa Wallner, Karl-Heinz T. Bäuml*

Department of Experimental Psychology, Regensburg University, Germany

ARTICLE INFO

Keywords:
Episodic memory
Retrieval
Cuing
Part-list cuing
Prose material

ABSTRACT

Numerous studies have shown that the presentation of some studied items as retrieval cues at test can impair recall of the remaining items. This effect, often referred to as part-list cuing (PLC) impairment, has mostly been demonstrated with simple word lists and short retention intervals between study and test. Across 4 experiments, this study examined the effects of PLC with educationally relevant prose passages for retention intervals of up to one week. Results showed that the effects of PLC with prose material depend critically on retention interval and test format. In the absence of any further retrieval cues, Experiments 1–3 found detrimental effects of PLC after short delay but neutral effects of PLC after delays of 2 days or 1 week. In the presence of gapped sentences ("fill-in-the-blank") serving as (additional) retrieval cues at test, Experiment 4 found a neutral effect of PLC after short delay but a beneficial effect after a delay of 2 days. With prose material, detrimental effects of PLC may thus be restricted to short retention interval and neutral or even beneficial effects may arise after prolonged retention interval. The findings suggest that both detrimental mechanisms — like blocking and inhibition – and beneficial mechanisms — like context reactivation – contribute to the effects of PLC with prose material.

1. Introduction

Retrieval cues can impair recall performance. This ironic finding is a core element of research on part-list cuing [PLC]. Since Slamecka's (1968) seminal work, research on PLC has investigated how the presentation of a random selection of studied items as retrieval cues at test influences recall of the remaining (target) items. The surprising result of this research has been that PLC often attenuates target recall (for an early review, see Nickerson, 1984). While indeed many studies on PLC have reported detrimental effects of cuing (e.g., Brown, 1968; Roediger, 1973; Sloman, 1991; Watkins, 1975), there is also evidence that PLC is not always detrimental. For instance, whereas PLC has often been found to be detrimental when employing short retention intervals of few minutes between study and test, in several recent studies the retention interval was prolonged up to several hours or even days and PLC was found to leave recall unaffected or to even increase recall levels (e.g., Bäuml & Schlichting, 2014; Lehmer & Bäuml, 2018a). Thus, in general, both detrimental and beneficial processes seem to contribute to the effects of PLC (for a more recent review, see Lehmer & Bäuml, 2018b).

Most work on PLC has been conducted in the verbal learning tradition using simple word lists as study material. The findings from this work, therefore, leave it open how PLC influences recall of more complex study material, like, for instance, educationally relevant prose passages. Indeed, whether results for word lists generalize to prose passages is unclear. Prose material may be encoded in a more integrated way than is typical for word lists and thus, for instance, lead to reduced levels of interference between the single memory contents, which may attenuate possible detrimental effects of PLC (e.g., Bäuml & Schlichting, 2014; see too below). To date only two studies examined effects of PLC with prose passages and provided mixed results. Whereas in the one study, neutral and even beneficial effects of PLC were reported (Bäuml & Schlichting, 2014), in the other study, a detrimental effect of PLC arose (Fritz & Morris, 2015).

The present study aimed to investigate the effects of PLC under educationally relevant conditions, using prose passages as material and retention intervals of up to a week. The experiments were conducted to shed light on whether (i) the effects of PLC with prose material vary with the length of the retention interval, (ii) both detrimental and beneficial processes contribute to the effects of PLC with this type of material, and (iii) more generally, PLC influences recall of prose passages similar to how it influences the recall of words. The results of the experiments will thus inform theories of PLC about whether prose material creates effects of PLC that go beyond those reported for word lists and are not covered by current accounts of PLC.

* Corresponding author at: Department of Experimental Psychology, Regensburg University, 93040 Regensburg, Germany.
E-mail address: karl-heinz.baeuml@ur.de (K.-H.T. Bäuml).

https://doi.org/10.1016/j.cognition.2020.104427
Received 29 May 2020; Received in revised form 28 July 2020; Accepted 6 August 2020
Available online 25 August 2020

1.1. Prior PLC work with word lists

Prior research on PLC has mostly employed word lists as material and a short retention interval of few minutes between study and test. Under these conditions, the effect of PLC turned out to be detrimental in most cases with only few exceptions to this "rule" (e.g., Aslan & Bäuml, 2007; Basden & Basden, 1995). These detrimental effects have been attributed to blocking, inhibition, or strategy disruption processes. The blocking and inhibition accounts assume that PLC induces covert retrieval of the cue items during recall, which then blocks or actively inhibits recall of the remaining target items (Bäuml & Aslan, 2004; Roediger, 1973; Rundus, 1973). The strategy disruption account claims that subjects create serial retrieval plans during study and the presentation of part-list cues at test then disrupts these strategies, which can impair recall performance (Basden et al., 1977; Basden & Basden, 1995). Importantly, the three processes are not mutually exclusive and may operate in different experimental conditions (see Bäuml & Aslan, 2006; Lehmer & Bäuml, 2018a).

However, the detrimental effects of PLC observed after short retention interval may not generalize to prolonged retention interval. Two recent studies provided corresponding evidence by varying the length of the retention interval between study and test. In the one study, PLC was found to be detrimental after a retention interval of 5 min but to be beneficial after a retention interval of 2 days (Bäuml & Schlichting, 2014, Experiment 1). In the other study, PLC was found to be detrimental after a retention interval of 4 min, but to be neutral or beneficial after retention intervals of 30 min or 24 h (Lehmer & Bäuml, 2018a). The observed beneficial effect as well as the neutral effect were attributed to the contribution of context reactivation processes to the effects of PLC. This explanation assumes that when context at test differs from context during study – which typically is the case after prolonged retention interval because context changes over time (Bower, 1972; Estes, 1955) – the part-list cues help reinstating the study context, so that context can act as a powerful retrieval cue for recall of the target items and thus increase recall levels (e.g., Bäuml & Samenieh, 2012). Thus, as it seems, both detrimental processes – like blocking, inhibition, and strategy disruption – and beneficial processes – like context reactivation – can contribute to the effects of PLC in recall of simple word lists (for details, see Lehmer & Bäuml, 2018b).

1.2. Prior PLC work with prose passages

Research on PLC with prose passages is rare and, to our knowledge, there are only two studies to date that addressed the issue. In the one study (Fritz & Morris, 2015, Experiment 2), subjects studied two texts, each of about 200–300 words in length (The History and Uses of Garlic and Neandertals), and, after reading through a text and a retention interval of about 6 min, were provided with half of the statements of the text to recall the text's remaining (target) statements (*PLC condition*), or recalled the text in the absence of any part-list cues (*no-PLC condition*). Results revealed a detrimental effect of PLC on recall of the target information. In the other study (Bäuml & Schlichting, 2014, Experiment 2), subjects studied two texts, each of about 1500 words in length (The Shaolin Temple and The Big Bang Theory), and, after study of a text, were provided gapped sentences from the text and asked to fill in the correct (target) item from the previously studied text (e. g., "The word Shaolin means young ____." [Answer: forest]). For half of the subjects, 10 sentences from the original text were provided intact to serve as (additional) part-list cues at test (*PLC condition*), whereas the other half of the subjects recalled the target information in the absence of the 10 sentences (*no-PLC condition*). Recall was measured after a 4-min and a 2-days retention interval. Results showed a neutral effect of PLC after the short delay but a beneficial effect of PLC after the long delay.

Although the results from the two studies suggest that, similar to word lists, both detrimental and beneficial processes may contribute to the effects of PLC with prose passages, the results provide no solid basis yet to draw strong conclusions on how PLC influences recall of prose passages. Partly, this is due to the fact that the results of the studies disagree in the observed effects, as is reflected by the finding of detrimental effects of PLC (Fritz & Morris, 2015) versus neutral effects of PLC (Bäuml & Schlichting, 2014) after short retention interval. Partly, it is due to the fact that, for prolonged retention interval, there are currently results from a single experiment only and replication and generalization of the finding is required before drawing more firm conclusions. Thus, a more thorough investigation of the effects of PLC with prose material is necessary. The present study addresses the issue.

The results of 4 experiments are reported in each of which subjects studied one of two prose passages and later were asked to recall (parts of) the passage. Subjects should recall the passages in the presence of several sentences of a passage serving as part-list cues (*PLC condition*), or in the absence of such sentences (*no-PLC condition*). In each experiment, at least two retention interval conditions were included, a short delay of few minutes and a longer delay of two days or one week. In addition, the experiments varied (i) material, (ii) how many sentences were provided as part-list cues at test, (iii) whether the sentences were presented in a random order or in a serial order that matched the order of the sentences in the original text, and (iv) whether – in both the PLC and the no-PLC conditions – gapped sentences from the text were provided to recall target information ("fill-in-the-blank"), or "free" recall without such sentences was employed at test. Indeed, employed test format may influence the effects of PLC and, for instance, underlie the difference in PLC effects as they were observed in the studies by Bäuml and Schlichting (2014) and Fritz and Morris (2015). The results of the 4 experiments will provide a fairly rich picture of the effects of PLC on recall of prose passages.

2. Experiment 1

The aim of Experiment 1 was to examine the effects of PLC using prose passage as study material and different delay intervals between study and test. Subjects studied a single prose passage (the text "Sea Otters") and, after a delay of 5 min, 2 days, or 1 week, were asked to recall the text material. Half of the subjects were provided 10 randomly selected sentences of the text as retrieval cues for recall of the remaining text material (*PLC condition*), whereas the other half of the subjects recalled the text in the absence of any retrieval cues (*no-PLC condition*). Using similar testing conditions, Fritz and Morris (2015, Experiment 2) recently examined the effects of PLC with other prose material and a retention interval of 6 min, and reported a detrimental effect of PLC. On the basis of this finding, one may thus expect a detrimental effect of PLC for the short 5-min delay condition of the present experiment as well.[1] Fritz and Morris (2015) did not include a long-delay condition. However, if the effects of PLC with prose material were similar to the effects of PLC with word lists – as might be speculated on the basis of Fritz and Morris' result in the short-delay condition –, PLC with prose material might show neutral or even beneficial effects of PLC after long delay, given that neutral or beneficial effects of PLC have been reported with word lists for several long retention interval conditions (Bäuml & Schlichting, 2014; Lehmer & Bäuml, 2018a). If so, PLC might induce a detrimental effect in the 5-min delay condition but not in the 2-days and 1-week delay conditions of Experiment 1.

[1] Bäuml and Schlichting (2014, Experiment 2) also employed prose material. However, in contrast to Fritz and Morris (2015) and present Experiments 1–3, gapped sentences were provided as (additional) retrieval cues for recall of the target information. We will refer to this study in more detail in Experiment 4 of the present study.

2.1. Method

2.1.1. Participants

240 students of Regensburg University participated in the experiment (M = 21.77 years, range = 18–32 years, 68.8% female and 31.3% male). They were equally distributed across the six between-subjects conditions, resulting in n = 40 subjects in each condition. We used previous PLC studies as a starting point for determining sample size, including studies with word lists as well as the two studies with prose material (e.g., Bäuml & Aslan, 2006; Bäuml & Schlichting, 2014; Fritz & Morris, 2015; Lehmer & Bäuml, 2018a). All subjects spoke German as native language and received monetary reward or course credit for participation.

2.1.2. Materials

We employed the text "Sea Otters", which already served as study material in Roediger and Karpicke's (2006) classic study on the testing effect. The passage had been selected from the reading comprehension section of a test-preparation book for the Test of English as a Foreign Language (TOEFL; Rogers, 2001). The translated passage was 270 words in length. Following Roediger and Karpicke (2006), the text was divided into 30 idea units for scoring purposes. We randomly selected four different sets of idea units, each with 10 units, to serve as part-list cues at test for a quarter of the subjects, respectively. In each case, the remaining 20 idea units served as target material.

2.1.3. Design

The experiment had a 2 × 3 design with the between-subjects factors of CUING (no PLC, PLC) and DELAY (5 min, 2 days, 1 week). In the no-PLC condition, participants were instructed to freely recall as many facts of the text as possible, whereas participants in the PLC condition were provided with 10 idea units of the text to serve as retrieval cues before they were asked to recall the remaining material. Delay conditions differed in the time passing between study and test phase.

2.1.4. Procedure

The study phase consisted of two 5-min periods during which participants were asked to read the prose passage (first period) or restudy the prose passage (second period). Subjects solved simple arithmetic problems for 2 min between the two periods and for another 5 min after the end of the second period. Participants in the long delay conditions were dismissed at this point and asked to come back 2 days or 1 week later. For the participants in the short delay condition, the test phase followed. At test, half of the participants were provided with 10 randomly selected facts of the text which were presented on a sheet of paper in a random order. Subjects were asked to read each of the sentences aloud. Sheets remained present during the whole testing phase. Subjects then had 10 min to recall the remaining facts of the text and write them down on a blank sheet of paper. Subjects were encouraged to use the provided sentences as retrieval cues for recall of the remaining material. The other half of the participants recalled the text in the absence of any retrieval cues.

2.2. Results

Following Roediger and Karpicke (2006), subjects' responses were scored by giving one point for each correctly recalled idea unit (out of 30).[2] Consistent with prior PLC work, we restricted analysis to target units.

Fig. 1 shows mean recall rates for the target units as a function of cuing and delay.[3] A 2 × 3 analysis of variance (ANOVA) with the between-subjects factors of CUING (no PLC, PLC) and DELAY (5 min, 2 days, 1 week) showed a significant main effect of CUING, $F(1,234) = 9.07$, $MSE = 131.51$, $p = .003$, $\eta^2 = 0.04$, depicting the overall detrimental effect of PLC, and a main effect of DELAY, $F(2,234) = 97.54$, $MSE = 131.52$, $p < .001$, $\eta^2 = 0.46$, indicating typical time-dependent forgetting. There was also a significant interaction between the two factors, $F(2,234) = 5.66$, $MSE = 131.51$, $p = .004$, $\eta^2 = 0.05$, suggesting that PLC affected target recall differently across delay conditions. Follow-up comparisons indeed revealed that PLC impaired target recall in the 5-min delay condition (44.50% vs. 33.00%), $t(78) = 4.34$, $p < .001$, $d = 0.97$, but did not impair target recall in the 2-days delay condition (23.88% vs. 23.00%) and the 1-week delay condition (14.13% vs. 13.13%), both $ts(78) < 1$, both $ds < 0.09$, The detrimental effect of PLC thus was restricted to the short delay condition.

There were four different sets of target and cue units in this experiment, which were counterbalanced across participants. Set did not influence results, however, as is indicated by a 2 × 3 × 4 ANOVA with the between-subjects factors of CUING, DELAY, and SET, which revealed no main effect of SET and no interaction of SET with any of the other factors, all $Fs < 2.55$, $MSEs = 129.13$, $ps > .056$, $\eta^2 s < 0.03$.

2.3. Discussion

The results showed a detrimental effect of PLC after the short retention interval but neutral effects of PLC after the two long retention intervals, indicating an influence of delay between study and test on the effects of PLC. The finding in the short delay condition is consistent with the previous result by Fritz and Morris (2015, Experiment 2), who also used prose material and observed a detrimental effect of PLC after short retention interval. The results in the long delay condition go beyond the prior work, demonstrating that the detrimental effect of PLC with prose material may be restricted to short delay conditions.

3. Experiment 2

A first aim of Experiment 2 was to replicate the basic finding of Experiment 1 with another prose passage. In addition, the experiment examined the possible role of number of part-list cues for recall of target information and included a two-stage testing procedure rather than a single recall test. Subjects studied the text "The Sun" and, after a delay of 5 min or 1 week, were asked to recall the text material in a first, critical test. In this test, one third of the subjects were cued with 8 randomly selected sentences of the text and another third with 16 randomly selected sentences; the rest of the subjects recalled the text in the absence of any part-list cues. On the second, final test, all subjects were asked to recall as many facts from the text as possible without any retrieval cues.

Several previous PLC studies with word lists employed a two-stage testing procedure, in which, in the PLC condition(s), part-list cues were present on the first recall test but were removed on the subsequent second test (e.g., Basden et al., 1977; Basden & Basden, 1995; Bäuml & Aslan, 2006; Lehmer & Bäuml, 2018a). According to the strategy disruption account of PLC, subjects build serial retrieval plans during study and the presentation of part-list cues at test then disrupts subjects' strategies, creating the often observed detrimental effect of PLC.

[2] Roediger III and Karpicke (2006) employed a rather strict scoring method, in which, to be scored correct, the recalled facts had to match with the wording in the text. We additionally employed a more liberal scoring method, in which each single fact was already scored correct if the gist of the fact was remembered regardless of whether the wording matched. Although recall rates were generally higher with liberal scoring, the pattern of results turned out equivalent for the two scoring methods. Therefore, results for the strict scoring method are reported only. The same picture arose for Experiments 2–4 to be reported below.

[3] The datasets of Experiments 1–4 are also available via https://osf.io/3yf6n/.

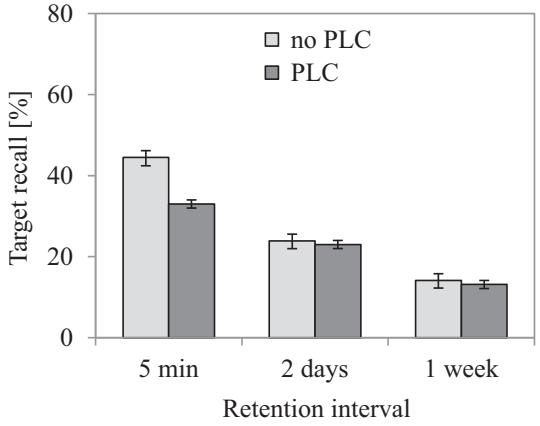

Fig. 1. Results of Experiment 1. Percentage of recalled target items is shown as a function of delay (5 min, 2 days, 1 week) and cuing (no PLC, PLC). Error bars represent standard errors.

Following Basden and Basden (1995), such disruption may disappear on a second test, when the part-list cues are removed and subjects are now free to recall the material in any order they like. In such case, the detrimental effect observed on the first, critical test may be eliminated on the second, final test. The strategy disruption account can explain PLC effects with word lists in certain experimental situations (e.g., Basden & Basden, 1995; Bäuml & Aslan, 2006). Using this two-stage testing procedure, we examined in Experiment 2 whether strategy disruption can also explain the detrimental effect of PLC with prose material, as it was observed in the short-delay condition of Experiment 1.

3.1. Method

3.1.1. Participants

Another 192 students of Regensburg University took part in the experiment (M = 21.82 years, range: 18–31 years, 76.6% female, 23.4% male). They were equally distributed across the six between-subjects conditions, resulting in n = 32 subjects in each condition. Compared to Experiment 1, sample size was slightly reduced based on the results of an analysis of test power conducted with the G*Power program (version 3, Faul et al., 2007). For this analysis, we set alpha = 0.05, beta = 0.20, and, motivated by the results of Experiment 1 above, effect size η^2 = 0.05 for the interaction between cuing and delay. Again, all subjects spoke German as native language and received monetary reward or course credit for participation.

3.1.2. Materials

We employed the text “The Sun”, which already served as study material in Roediger and Karpicke's (2006) testing effect study. Like the text “Sea Otters”, the passage had been selected from the reading comprehension section of a test-preparation book for the Test of English as a Foreign Language (Rogers, 2001). The translated passage was 251 words in length. For scoring purposes, the same 30 idea units were defined as in Roediger and Karpicke's study. From these 30 units, we randomly selected four different sets of idea units, each with 14 units, to serve as target material for a quarter of the subjects, respectively. In each case, the remaining 16 units served as part-list cues in the 16-part-list-cues condition, and a random selection of 8 units out of the 16 units served as part-list cues in the 8-part-list-cues condition.

3.1.3. Design

The experiment had a 3 × 2 design with the between-subjects factors of CUING (0 part-list cues, 8 part-list cues, 16 part-list cues) and DELAY (5 min, 1 week). In the 0-part-list-cues condition, participants were instructed to freely recall as many facts of the text as possible, while participants in the 8-part-list-cues and 16-part-list-cues conditions were provided with 8 respectively 16 units of the text to serve as retrieval cues at test, before they were asked to recall the remaining facts of the text. Delay conditions differed in the time passing between study and test phases.

3.1.4. Procedure

The procedure followed Experiment 1 with four modifications. First, subjects studied the text “The Sun” rather than the text “Sea Otters”. Second, we omitted the 2-days delay condition given that in Experiment 1 it led to the same results as the 1-week delay condition. Third, we varied the number of part-list cues that were provided at test; while in Experiment 1, subjects in the PLC condition were cued with 10 idea units, in Experiment 2, subjects were cued with 8 or 16 idea units. Finally, the test phase was changed to a two-stage testing procedure. In the first, critical test part-list cues were provided in the 8-part-list-cues and 16-part-list-cues conditions – but not in the 0-part-list-cues condition –, whereas on the second, final test, in none of the three conditions any part-list cues were provided. Between the two tests, all subjects were asked to count backwards from a three-digit number for 30 s and to work on the standard progressive matrices (Raven, 1998) for another 4 min.

On the critical test, after reading the cues aloud, subjects in the 16-part-list-cues condition had 5 min to freely recall the remaining units of the text (maximum: 14 target units); subjects in the 8-part-list-cues condition had 7.5 min to recall the remaining units (maximum: 22 units = 14 target units + 8 nontarget units); and subjects in the 0-part-list-cues condition had 10 min to recall all text details (maximum: 30 units = 14 target units + 16 nontarget units). This procedure contrasts with the one employed in Experiment 1, in which the duration of the recall period on the (critical) test did not vary with PLC conditions, and subjects in the no-PLC condition recalled many more items (target plus nontarget items) than in the PLC condition within the same time interval. The change in duration of recall period from Experiment 1 to Experiment 2, did not influence results, however, as can be seen in Section 3.2 below. The final test was identical for all subjects. Each subject had 10 min to recall as many of the facts of the text as they could remember and to write them down on a blank sheet of paper.

3.2. Results

Scoring of subjects' recall responses again followed Roediger and Karpicke (2006). The responses were scored by giving one point for each correctly recalled idea unit (out of 30).

Fig. 2 shows mean recall rates for the target units, in both the critical test (A) and the final test (B). A 3 × 2 × 2 ANOVA with the between-subjects factors of CUING (0 part-list cues, 8 part-list cues, 16 part-list cues) and DELAY (5 min, 1 week) and the within-subject factor of TEST (critical test, final test) showed no main effect of TEST, F

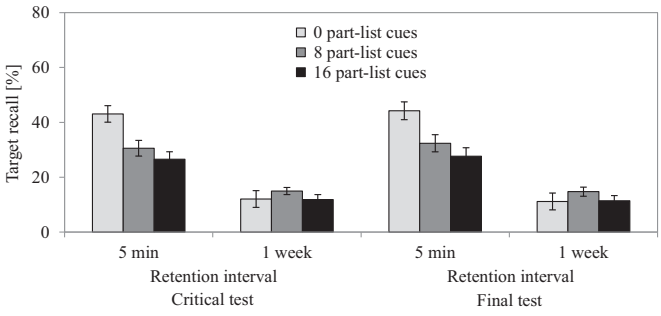

Fig. 2. Results of Experiment 2. Percentage of recalled target items is shown as a function of delay (5 min, 1 week) and cuing (0 part-list cues, 8 part-list cues, 16 part-list cues), separately for the critical test and the final test. Error bars represent standard errors.

(1,186) < 1, and no interaction of TEST with any of the other factors, all *Fs* < 3.40, *MSEs* = 24.52, *ps* > .067, η^2*s* < 0.02. However, there was a significant main effect of CUING, *F*(2,186) = 6.07, *MSE* = 360.68, *p* = .003, η^2 = 0.06, depicting the overall detrimental effect of PLC, a significant main effect of DELAY, *F*(1,186) = 121.82, *MSE* = 360.68, *p* < .001, η^2 = 0.40, indicating time-dependent forgetting, and a significant interaction between CUING and DELAY, *F*(2,186) = 7.57, *MSE* = 360.68, *p* = .001, η^2 = 0.08, suggesting that PLC affected recall differently in the two delay conditions.

Separate follow-up analyses for the two delay conditions revealed a significant main effect of CUING in the short delay condition, *F*(2, 93) = 8.67, *MSE* = 541.37, *p* < .001, η^2 = 0.16. PLC impaired target recall both when 8 part-list cues were provided at test (43.64% vs. 31.47%), *F*(1, 62) = 8.56, *p* = .005, *MSE* = 553.76, η^2 = 0.12, and when 16 part-list cues were provided (43.64% vs. 27.12%), *F*(1, 62) = 16.01, *p* < .001, *MSE* = 545.88, η^2 = 0.21. There was no difference in recall rates between the two PLC conditions, however, *F*(1, 62) = 1.16, *p* = .286, *MSE* = 524.44, η^2 = 0.02. In contrast to the short-delay condition, there was no significant effect of cuing in the long-delay condition (11.61% vs. 14.85% vs. 11.61%), *F*(2, 93) = 1.24, *MSE* = 180.00, *p* = .294, η^2 = 0.03.

Two further analyses were conducted. First, like in Experiment 1, there were four different sets of target and cue units, which were counterbalanced across participants. However, for both the critical and the final test, a 3 × 2 × 4 ANOVA with the between-subjects factors of CUING, DELAY, and SET revealed no main effect of SET and no interaction of SET with any of the other factors, all *Fs* < 1.53, *ps* > .209. Second, to exclude that the neutral effect of PLC after long delay was due to a floor effect, we analyzed whether results in this condition would change when data from the liberal rather than the strict scoring method were employed (compare Footnote 2 above). The corresponding 3 × 2 × 2 ANOVA with the between-subjects factor of CUING and the within-subject factors of TEST and SCORING (strict, liberal) showed a main effect of scoring (12.07% vs. 27.05%), *F*(1,93) = 166.80, *p* < .001, *MSE* = 118.69, η^2 = 0.64, but no interaction of scoring with the other two factors, both *Fs* < 1, indicating that results in the long delay condition were not due to a floor effect.

3.3. Discussion

Using different material than employed in Experiment 1, the results of Experiment 2 replicate the basic finding of Experiment 1, with a detrimental effect of PLC after short delay and a neutral effect of PLC after long delay. Going beyond Experiment 1, the results of Experiment 2 show no influence of number of provided part-list cues on the effects of PLC, and no difference in PLC effects between a first, critical test and a second, final test. Prior work with word lists reported an influence of number of provided part-list cues on the detrimental effect of PLC, with a higher amount of recall impairment if number of part-list cues was increased (Roediger, 1973; Roediger et al., 1977; see also Dagnall et al., 2008). The present results for the short delay condition show a similar numerical trend but no statistical difference between the 8- and 16-part-list-cues conditions.[4] One reason for this discrepancy may be that number of part-list cues relative to number of employed target items was higher in the previous studies than in present Experiment 2 – for instance, 16 and 32 cues items out of 48 study items in Roediger et al. (1977) versus 8 and 16 cue items out of 30 study items in the present study –, which may have led to generally larger detrimental effects of PLC in the previous studies and an increased likelihood to detect differences between the single cuing conditions. Future work may therefore revisit the issue employing higher numbers of part-list cues than was done here.

The finding that the detrimental effect of PLC in the short delay condition was equally present in the critical and final recall tests suggests that strategy disruption did not underlie the detrimental effects reported in Experiments 1 and 2. According to the strategy disruption account, part-list cues disrupt subjects' preferred output orders at test, which should create recall impairment on the critical test, where cues are provided, but not on the final test, where the cues are removed (e.g., Basden & Basden, 1995; Bäuml & Aslan, 2006). Rather, the results appear more consistent with blocking and inhibition views of the effect. Indeed, both blocking and inhibition should create lasting PLC effects, which should not disappear on a final test in which the cues are removed (see Aslan et al., 2007; Bäuml & Aslan, 2006). The results of Experiment 2 thus favor blocking and inhibition views of the observed detrimental effects over a strategy disruption view (see also General discussion section).

4. Experiment 3

There were two major goals with Experiment 3. The one goal was to replicate the finding of Experiment 2 that the effects of PLC do not change from a first, critical test to a second, final test with other prose material. We chose the material employed in Experiment 1. The second goal was to examine whether the serial order of the part-list cues provided at test influences the effects of PLC. We therefore replaced the 16-part-list-cues condition used in Experiment 2 by another 8-part-list-cues condition, in which the cues were not presented in a random order but were presented in a serial order that matched the cues' order in the studied text. The rest of the experiment was identical to Experiment 2 with the only other difference that a 2-days rather than a 1-week retention interval was employed in the long delay condition.

The strategy disruption account of PLC suggests that, after short delay, PLC may be much less disruptive, or even nondetrimental, if the order of the items provided as part-list cues at test matched the order of the items as they were presented during study. Indeed, if subjects created serial retrieval plans during study and the provided part-list cues followed the serial order of the text, then possible effects of strategy disruption should be attenuated and the recall impairment be reduced, if not eliminated. The results of studies with word lists support this prediction for certain experimental conditions (e.g., Basden & Basden, 1995; but see Slamecka, 1969). Using prose passages as material, Fritz and Morris (2015) addressed the issue in a between-experiments comparison, in which part-list cues were presented in a random order in one experiment (Experiment 2) and in the correct serial order in the other (Experiment 1). They did not find a reduction of the detrimental effect of PLC when the serial order of the cues was maintained and even found an increase of the detrimental effect, relative to the random order condition. Experiment 3 addressed the issue by comparing the two cuing conditions within a single experiment, in both a short-delay and a long-delay condition.

4.1. Method

4.1.1. Participants

Another 192 students of Regensburg University took part in the experiment (*M* = 22.65 years, range = 18–36 years, 79.2% female). Sample size followed Experiment 2. Again subjects were equally distributed across the six between-subjects conditions, resulting in *n* = 32 subjects in each condition. All subjects spoke German as native language and received monetary reward or course credit for their participation.

4.1.2. Materials

We employed the same prose passage as already used in Experiment 1, i.e., the “Sea Otters” text. Out of the text's 30 idea units we randomly

[4] We also analyzed the nonsignificant difference between the two PLC conditions by running Bayes factor analysis (Lee & Wagenmakers, 2013). The resulting Bayes factor of 0.21 indicates moderate evidence in favor of the null hypothesis.

selected four different sets of 22 idea units to serve as target material for a quarter of the subjects, respectively. In each case, the remaining 8 idea units served as part-list cues in this experiment.

4.1.3. Design

The experiment had a 3 × 2 design with the between-subjects factors of CUING (no PLC, PLC random order, PLC serial order) and DELAY (5 min, 2 days). In the no-PLC condition, participants were instructed to freely recall as many facts of the text as possible, whereas participants in the two PLC conditions were provided with 8 units of the text before they were asked to recall the remaining material. In the PLC-random-order condition, the order of the cues was random, while in the PLC-serial-order condition, the order of the cues matched the cues' order in the text. The two delay conditions differed in the time passing between study and test.

4.1.4. Procedure

The procedure was largely identical to Experiment 2, with three changes only. The first change was that the 16-part-list-cues condition employed in Experiment 2 was replaced by the PLC-serial-order condition. Second, the 1-week delay condition was replaced by a 2-days delay condition. Third, like in Experiment 1, the recall period was 10 min in all experimental conditions.

4.2. Results

Fig. 3 shows mean recall rates for the target units, for both the critical test (A) and the final test (B). A 3 × 2 × 2 ANOVA with the between-subjects factors of CUING (no PLC, PLC random order, PLC serial order) and DELAY (5 min, 2 days) and the within-subject factor of TEST (critical test, final test) showed no main effect of TEST, $F(1,186) = 1.73$, $MSE = 26.11$, $p = .190$, $\eta^2 = .01$, and no interaction of TEST with any of the other factors, all $Fs < 3.81$, $MSEs = 26.11$, $ps > .052$, $\eta^2 s < 0.02$. However, again there were a significant main effect of CUING, $F(2,186) = 3.36$, $MSE = 384.77$, $p = .037$, $\eta^2 = 0.04$, depicting the overall detrimental effect of PLC, a significant main effect of DELAY, $F(1,186) = 21.38$, $MSE = 384.78$, $p < .001$, $\eta^2 = 0.10$, reflecting time-dependent forgetting, and a significant interaction between the two factors, $F(2,186) = 6.74$, $MSE = 384.77$, $p = .001$, $\eta^2 = 0.07$, indicating that the effects of cuing varied with delay.

Separate follow-up analyses for the two delay conditions showed a significant main effect of cuing in the short delay condition, $F(2,93) = 8.47$, $MSE = 426.11$, $p < .001$, $\eta^2 = 0.15$. PLC impaired target recall both when the cues were provided in a random order (37.75% vs. 29.19%), $F(1,62) = 4.38$, $p = .040$, $MSE = 513.23$, $\eta^2 = 0.07$, and when they were provided in the correct serial order (37.75% vs. 22.59%), $F(1,62) = 17.62$, $p < .001$, $MSE = 407.87$, $\eta^2 = 0.22$. There was no statistical difference in recall rates between the two PLC conditions, $F(1,62) = 3.91$, $p = .053$, $MSE = 357.25$, $\eta^2 = 0.06$. In contrast to the short-delay condition, there was no significant effect of PLC in the long-delay condition (18.33% vs. 22.44% vs. 20.81%), $F(2,93) < 1$, $\eta^2 = 0.02$.

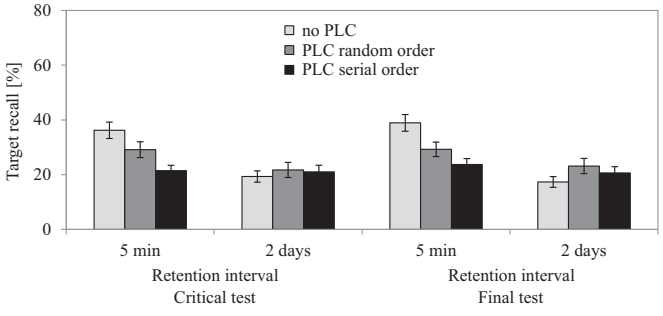

Fig. 3. Results of Experiment 3. Percentage of recalled target items is shown as a function of delay (5 min, 2 days) and cuing (no PLC, PLC serial order, PLC random order), separately for the critical test and the final test. Error bars represent standard errors.

Like in Experiments 1 and 2, there were four different sets of target and cue units, which were counterbalanced across participants. Again, for both the critical and the final test, a 3 × 2 × 4 ANOVA with the between-subjects factors of CUING, DELAY, and SET revealed no main effect of SET and no interaction of SET with any of the other factors, all $Fs < 2.16$, $ps > .094$.

4.3. Discussion

The results of Experiment 3 provide another demonstration that PLC with prose material induces a detrimental effect after short delay but a neutral effect after long delay. The results also replicate the finding of Experiment 2 that the pattern of PLC effects does not change from a first, critical test to a second, final test. In particular, the results show that the size of the detrimental effect of PLC is not reduced when the random order of the part-list cues is replaced by the correct serial order of the cues, i.e., the order in which the cue sentences were provided in the text. There was even a numerical trend towards a larger detrimental effect when the cues were provided in the correct serial order, which mimics Fritz and Morris' (2015) result in their between-experiments comparison.[5] The finding thus challenges a strategy disruption view of the detrimental effect, according to which PLC would be expected to be less disruptive if the order of the items provided as part-list cues at test matched the order of the items during study. Importantly, there was also no difference between the serial-order and the random-order conditions after long delay, suggesting that, regardless of delay, order of part-list cues at test does not have a major influence on the effects of PLC with prose material.

5. Experiment 4

All experiments reported above found a detrimental effect of PLC after short delay and a neutral effect of PLC after long delay. This pattern deviates from Bäuml and Schlichting's (2014, Experiment 2) finding. Comparing PLC effects between a short 5-min and a long 2-days retention interval, these researchers reported a neutral effect of PLC after the short delay and a beneficial effect of PLC after the long delay. There are several methodological differences between this study and the present experiments, but two differences stand out. The one difference is the length of the employed prose passages. Whereas Bäuml and Schlichting employed passages with about 1500 words in length, the present experiments employed passages with less than 300 words in length. The other difference is the employed test format. Indeed, while in the present experiments, apart from part-list cues, subjects recalled the target information in the absence of any retrieval cues, in Bäuml and Schlichting's experiment, gapped sentences were provided at test in both the PLC and the no-PLC conditions, and subjects were asked to fill in the correct (target) item from the previously studied text (e. g., "The Hubble telescope found the heavy element ____ in extremely ancient stars." [Answer: boron]; see also Chan et al., 2006).

Critically, the presentation of gapped sentences as retrieval cues at test may reduce or even eliminate interference from other information, which may then reduce, or even eliminate, possible effects of blocking and inhibition, as they may arise in response to PLC (Bäuml & Aslan, 2004; Roediger, 1973). As a result, the presence of gapped sentences at test may attenuate the detrimental effect of PLC after short delay. Moreover, if, in general, not only blocking and inhibition but also context reactivation processes operated in response to PLC, and context reactivation operated mainly after long delay (Bäuml & Schlichting,

[5] Bayes factor analysis revealed a Bayes factor of 2.48, indicating anecdotal evidence for the hypothesis of a difference between cuing conditions.

2014; Lehmer & Bäuml, 2018a), then, in the presence of gapped sentences, blocking and inhibition should hardly contribute to the effects of PLC and context reactivation thus induce an increase in recall levels in the long-delay condition. In the presence of gapped sentences, PLC may therefore lead to a neutral effect on recall after short delay and a beneficial effect after long delay.

We addressed the issue in Experiment 4. Experiment 4 was largely identical to Experiment 1 with the main difference that gapped sentences were provided as retrieval cues at test, in both the PLC and the no-PLC conditions. If the presence versus absence of gapped sentences at test underlied the difference in results between present Experiments 1–3 and Bäuml and Schlichting's (2014) experiment, then in Experiment 4, a neutral effect of PLC may arise after short delay and a beneficial effect after long delay. We examined the issue using the prose passage employed in Experiment 1.

5.1. Method

5.1.1. Participants

Another 128 students of Regensburg University participated in the experiment (M = 22.09 years, range = 18–34 years, 63.3% female). They were equally distributed across the four between-subjects conditions, resulting in n = 32 subjects in each condition. Sample size was determined by an analysis of test power conducted with the G*Power program, by setting alpha = 0.05, beta =0.20, and effect size η^2 = 0.05 for the interaction between cuing and delay (see Bäuml & Schlichting, 2014, Experiment 2). All subjects spoke German as native language. In exchange for participation, course credit or monetary reward was provided.

5.1.2. Materials

The same prose passage as employed in Experiment 1 ("Sea Otters") was used as material. From the 30 idea units contained in the text, we randomly selected four different sets of 10 units, each serving as part-list cues, and, from the remaining 20 units, selected another 10 units, each serving as target items for a quarter of the subjects, respectively. Each target item was tested in the form of a gapped sentence, like "Sea otters sleep often on masses of ___." [Answer: kelp] or "Sea otters are ___ long." [Answer: 4 to 5 ft]).

5.1.3. Design

The experiment had a 2 × 2 design with the between-subjects factors of CUING (no PLC, PLC) and DELAY (5 min, 2 days). Participants were provided gapped sentences for recall of the target information in the presence of part-list cues (PLC condition) or in their absence (no-PLC condition). Again, delay conditions differed in the time passing between study and test.

5.1.4. Procedure

The procedure followed the one used in Experiment 1, with two modifications. The one modification was that only one long delay condition (2 days) was included in the experiment. The other modification was that gapped sentences were provided at test for recall of the target items, in both the PLC and the no-PLC conditions. These sentences were presented successively and in a random order. Following Bäuml and Schlichting (2014), subjects had 25 s for each sentence to recall the target item. Responses were given orally and experimenters recorded the answers. After target recall, all subjects were asked to freely recall all remaining information from the text, but responses were not analyzed.

5.2. Results

Fig. 4 shows mean recall rates for the target items. A 2 × 2 ANOVA with the between-subjects factors of CUING (no PLC, PLC) and DELAY (5 min, 2 days) showed a main effect of DELAY, F(2,124) = 11.11, MSE = 425.83, p = .001, η^2 = 0.08, indicating typical time-dependent forgetting. There was no significant main effect of CUING, F(1,124) = 2.10, MSE = 425.83, p = .150, η^2 = 0.02, and no significant interaction between the two factors, F(2,124) = 2.10, MSE = 425.83, p = .150, η^2 = 0.02. However, planned comparisons revealed that, whereas PLC did not influence target recall in the 5-min delay condition (64.69% vs. 64.69%), t(62) < 1, there was a significant beneficial effect of PLC in the 2-days delay condition (47.25% vs. 57.81%), t(62) = 2.02, p = .047, d = 0.51.[6]

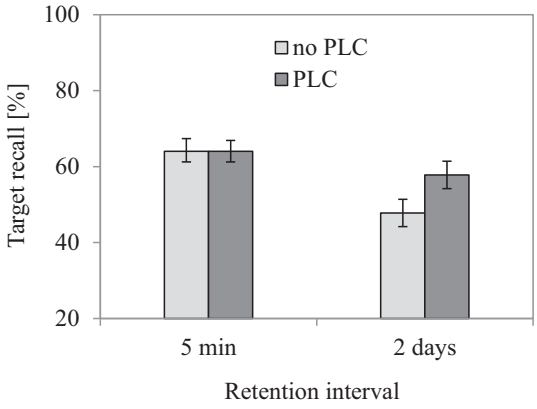

Fig. 4. Results of Experiment 4. Percentage of recalled target items is shown as a function of delay (5 min, 2 days) and cuing (no PLC, PLC). Error bars represent standard errors.

Next, we compared the effects of PLC between Experiment 1 (when gapped sentences were absent) and Experiment 4 (when gapped sentences were present) directly, separately for the 5-min and the 2-days delay conditions. Corresponding 2 × 2 ANOVAs with the between-subjects factors of CUING (no PLC, PLC) and EXPERIMENT (Experiment 1, Experiment 4) revealed a significant interaction between the two factors, both after the 5-min delay, F(1,140) = 4.39, MSE = 264.81, p = .038, η^2 = 0.03, and the 2-days delay, F(1,140) = 4.48, MSE = 262.67, p = .036, η^2 = 0.03. These interactions indicate that test format influenced the effects of PLC in both delay conditions.

5.3. Discussion

The results of Experiment 4, when compared to those of Experiment 1, show that the presence of gapped sentences as (additional) retrieval cues at test can turn (i) the detrimental effect of PLC after short delay into a neutral effect and (ii) the neutral effect of PLC after long delay into a beneficial effect, thus pointing to a critical role of test format for the effects of PLC with prose material, both after short delay and after long delay. The finding of a neutral effect of PLC after short delay and a beneficial effect of PLC after long delay in Experiment 4 mimics the one reported in Bäuml and Schlichting (2014, Experiment 2). Using other prose material, these researchers had reported a nonsignificant detrimental effect of 6.2% after short delay and a significant beneficial effect of 11.1% after long delay, which compares well with the present results of the perfectly neutral effect of PLC after short delay and the significant beneficial effect of 10.6% after long delay.

Bäuml and Schlichting (2014) attributed the observed neutral effect after short delay to reduced interference at test, arising from the fact that the presence of gapped sentences may largely attenuate interference from other information, and thus reduce possible effects of blocking and inhibition in response to PLC. In particular, they

[6] Like in Experiments 1–3, the different sets of target and nontarget items did not affect results. A corresponding ANOVA revealed no main effect of SET and no interaction of SET with any of the other factors, all Fs < 1.69, $MSEs$ = 433.07, ps > .173.

attributed the observed beneficial effect after long delay to context reactivation processes, assuming that when context at test differs from context during study – which should occur after longer delay (e.g., Estes, 1955) – the part-list cues help reinstating the study context, so that context can act as a powerful retrieval cue for recall of the target items. The results of Experiment 4 are consistent with these views.

The proposal that context reactivation processes mediate the beneficial effect of PLC after long retention interval can explain results for both prose passages and word lists (see Bäuml & Schlichting, 2014), and thus provides a parsimonious account of the effect across different study materials. This holds while, in principle, the beneficial effect for prose passages could also arise if the cues provided the participants with additional semantic information about the topic, which might enhance recall levels. However, such additional semantic information would be present after both short and long retention interval, and thus should lead to beneficial effects also after short retention interval, which is not what the present results show. Still, by varying the intrinsic meaning of the cues to the story, future work may address the possible contribution of semantic processes to the effects of PLC with prose passages.

6. General discussion

The results of Experiments 1–3 found the effects of PLC with prose material to be detrimental after a short retention interval of 5 min and to be neutral after prolonged retention intervals of 2 days and 1 week. This pattern arose for different materials, different numbers of part-list cues, different serial orders of part-list cues, and, using a two-stage testing procedure, in both a first, critical and a second, final recall test. Results for the short delay condition are consistent with the results of the previous study by Fritz and Morris (2015), who found detrimental effects of PLC after a short 6-min delay with prose material. Results for the two long delay conditions go beyond this previous study, in which the influence of long delay had not yet been examined.

The results of Experiment 4 show a different pattern than the results of Experiments 1–3, with a neutral effect of PLC after short delay and a beneficial effect after long delay. Critically, recall of the target information in Experiment 4 occurred in the presence of gapped sentences serving as (additional) retrieval cues, and thus differed from the test format employed in Experiments 1–3, in which no such retrieval cues were provided. The finding of Experiment 4 is consistent with the previous study by Bäuml and Schlichting (2014, Experiment 2), who used other prose material but the same testing procedure as employed in Experiment 4, and also found a neutral effect of PLC after short delay but a beneficial effect after prolonged delay. The findings from the present experiments together with those from the two previous studies thus suggest a critical role of test format for the effects of PLC with prose material, with neutral effects (short delay) and beneficial effects (long delay) in the presence of gapped sentences, and detrimental effects (short delay) and neutral effects (long delay) in their absence.

6.1. Cognitive mechanisms of PLC after short delay

When using word lists as study material, the typical detrimental effect of PLC after short delay has mostly been attributed to blocking, inhibition, or strategy disruption processes. According to the strategy disruption account, subjects try to develop individual retrieval plans during encoding based on their subjective organization of the list items. When in the PLC condition a randomly selected subset of items is presented in a random order at test, these retrieval cues then disrupt the preferred recall order and the participants switch to a less effective order, which reduces recall performance (Basden & Basden, 1995). On the basis of this account, the expectation arises that the detrimental effect should be present on a first, critical test, in which the part-list cues are provided, but be eliminated on a subsequent, final test, in which the part-list cues are removed. In addition, part-list cues that preserve the order in which the cue items were presented during study should be less detrimental compared to a condition in which the cues were presented in a random order (e.g., Basden & Basden, 1995). Results of both Experiment 2 and Experiment 3 disagree with these expectations. They show no reduction of the detrimental effect from a first, critical to a subsequent, final test, and no reduction of the detrimental effect when the part-list cues were presented in a serial order that matched the order of the cue items during study. The detrimental effects of PLC in the present study thus should not have been caused by strategy disruption.

Blocking and inhibition may provide better explanations of the detrimental effects as they were found in the present study. Blocking and inhibition assume that PLC induces covert retrieval of the cue items, which then block or actively inhibit recall of the remaining target items (Bäuml & Aslan, 2004; Roediger, 1973). According to these accounts, the detrimental effect of PLC should not be eliminated from a first, critical to the subsequent, final test, because the effects of both blocking and inhibition should be lasting (see Aslan et al., 2007; Bäuml & Aslan, 2006). The present finding of lasting PLC effects thus is consistent with the two accounts. Inhibition and blocking can also explain why the detrimental effect of PLC after short delay disappeared when gapped sentences were provided as retrieval cues in Experiment 4. Because this test format should reduce interference from other memory contents, PLC should not induce much blocking or inhibition and thus not much impair recall performance. Results in the short delay condition, therefore, favor inhibition and blocking explanations over a strategy disruption explanation of the observed results.[7]

6.2. Cognitive mechanisms of PLC after long delay

Two previous studies reported beneficial effects of PLC after prolonged retention interval, both when using word lists as material (Bäuml & Schlichting, 2014, Experiment 1; Lehmer & Bäuml, 2018a) and when using prose passages (Bäuml & Schlichting, 2014, Experiment 2). These beneficial effects were attributed to context reactivation processes. Following the general view that delay induces context drift and thus reduces the contextual overlap between study and test (e.g., Estes, 1955), the proposal has been that part-list cues reinstate the study context and the reinstated context then serves as a powerful retrieval cue for recall of the target items, thus improving recall performance (see too Bäuml & Samenieh, 2012).

The proposal that PLC-induced context reactivation processes operate after long delay is consistent with the beneficial effect of PLC observed in present Experiment 4 but, at first glance, may seem in conflict with the results in the long-delay conditions of Experiments 1–3, which showed neutral effects of PLC. However, the difference in results may reflect a difference in interference levels between experiments, with higher interference between memory contents in the absence of gapped sentences at test (Experiments 1–3) than in their presence (Experiment 4). On the basis of this view, context reactivation processes may have operated in all four experiments, but potential beneficial effects of context reactivation may have been masked by blocking and inhibition processes in Experiments 1–3, in which interference levels were high, but not in Experiment 4, in which interference levels were low. As a net result, PLC may have induced a neutral effect on recall in Experiments 1–3 but a beneficial effect in Experiment 4.

[7] Arguably, the finding that the serial order of the cue items at test had no major influence on the size of the detrimental effect is also in line with blocking and inhibition. In fact, assuming that serial order should not much influence strength of the single cue items, the cue items should block or inhibit recall of the target items more or less regardless of their exact serial order.

6.3. From word lists to prose passages

On the basis of the views provided in Sections 6.1 and 6.2 above, the present results suggest that similar cognitive processes may underlie the effects of PLC with prose material and word lists. Indeed, also with word lists, detrimental, neutral, and beneficial effects of PLC have been observed and, also with word lists, the effects have been attributed to different relative contributions of detrimental and beneficial processes on recall performance (e.g., Lehmer & Bäuml, 2018b). On the other hand, the results of Experiments 1–3 indicate that, after long delay, beneficial effects of PLC may not easily arise with prose material – at least in the absence of gapped sentences – whereas robust beneficial effects of PLC have been reported with word lists even in the absence of additional retrieval cues. Several factors may be responsible for the difference in findings. One such factor may be a difference in interference levels between employed prose passages and word lists. While the beneficial effects observed previously with word lists have been reported for relatively short word lists, the neutral effects observed here with prose passages occurred with relatively long prose passages. This difference in amount of studied material may have caused a difference in induced interference levels at test and may have led to higher relative contributions of blocking and inhibition with the prose passages than the word lists – which may have masked the beneficial effects of context reactivation for the prose material after long delay. Future work should address the issue in more detail and examine the possible role of amount of encoded information for the effects of PLC, with both word lists and prose passages.

The present results favor an explanation of the detrimental effect of PLC through blocking and inhibition processes over a strategy disruption explanation. However, research with word lists identified a role of strategy disruption for the detrimental effect of PLC, although mainly in situations in which subjects encoded the material through repeated study-test cycles or when subjects were explicitly encouraged to create serial retrieval plans during study (e.g., Basden & Basden, 1995; Bäuml & Aslan, 2006; see too Lehmer & Bäuml, 2018b). No such conditions were met in the present experiments, in which subjects encoded the material within successive study cycles. Future work should examine whether strategy disruption can play a role in PLC with prose passages as well, for instance, by letting subjects encode the passages through repeated study-test cycles. The results from such studies will provide further insights into which processes can influence the effects of PLC with prose passages.

7. Conclusions

Most previous studies on PLC reported detrimental effects of PLC, thus contributing to the general picture that PLC typically impairs recall performance. Together with the results of recent studies with word lists (Bäuml & Schlichting, 2014; Lehmer & Bäuml, 2018a), the present study with prose material challenges this picture, indicating that the finding of detrimental effects of PLC may be restricted to short retention intervals between study and test. Indeed, both the present study and the prior work found neutral or even beneficial effects of PLC when the retention interval was prolonged. In daily life, the delay between encoding and recall is often prolonged – like in educational settings when students are interrogated about the contents of a previous lecture, or in eyewitness testimony situations when witnesses are interviewed about the details of a crime – suggesting that, in daily life, PLC may often be neutral or even beneficial for recall performance. The findings open a new arena for PLC research and the investigation of the question of how PLC influences recall when recall is delayed.

Software

All experiments reported in this manuscript were implemented using the software PowerPoint 1909 (Microsoft Corporation). The software was run on standard desktop computers with the operating system Windows 7 (Microsoft, Redmond, WA). All data were analyzed using IBM SPSS Statistics for Windows, Version 25.0 (IBM Corp., Armonk, NY) and G*Power 3.1 (Faul et al., 2007).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Lisa Wallner: Methodology, Formal analysis, Writing - original draft. **Karl-Heinz T. Bäuml:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors have no competing interests to declare.

Appendix A. Supplementary material

The study materials employed in the present experiments as well as the data from the single experiments are available on the Open Science Framework (https://osf.io/3yf6n/).

References

Aslan, A., & Bäuml, K.-H. T. (2007). Part-list cuing with and without item-specific probes: The role of encoding. *Psychonomic Bulletin and Review, 14*, 489–494.
Aslan, A., Bäuml, K.-H. T., & Grundgeiger, T. (2007). The role of inhibitory processes in part-list cuing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 335–341.
Basden, D. R., & Basden, B. H. (1995). Some tests of the strategy disruption interpretation of part-list cuing inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1656–1669.
Basden, D. R., Basden, B. H., & Galloway, B. C. (1977). Inhibition with part-list cuing: Some tests of the item strength hypothesis. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 100–108.
Bäuml, K. H.-T., & Aslan, A. (2006). Part-list cuing can be transient and lasting: The role of encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 33–43.
Bäuml, K.-H. T., & Aslan, A. (2004). Part-list cuing as instructed retrieval inhibition. *Memory and Cognition, 32*, 610–617.
Bäuml, K.-H. T., & Samenieh, A. (2012). Infuences of part-list cuing on different forms of episodic forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 366–375.
Bäuml, K.-H. T., & Schlichting, A. (2014). Memory retrieval as a self-propagating process. *Cognition, 132*, 16–21.
Bower, G. H. (1972). A selective review of organizational factors in memory. In E. Tulving, & W. Donaldson (Eds.). *Organization of memory* (pp. 93–137). New York: Academic Press.
Brown, J. (1968). Reciprocal facilitation and impairment in free recall. *Psychonomic Science, 10*, 41–42.
Chan, J. C. K., McDermott, K. B., & Roediger, H. L., III (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135*, 553–571.
Dagnall, N., Parker, A., & Munley, G. (2008). The effects of part-set cuing on the recall of general knowledge information. *Journal of Cognitive Education and Psychology, 7*, 248–265.
Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62*, 145–154.
Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A exible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
Fritz, C. O., & Morris, P. E. (2015). Part-set cuing of texts, scenes, and matrices. *British Journal of Psychology, 106*, 1–21.
Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian modeling for cognitive science: A practical course.* Cambridge: University Press.
Lehmer, E.-M., & Bäuml, K.-H. T. (2018a). Part-list cuing can impair, improve, or not influence recall performance: The critical roles of encoding and access to study context at test. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 1186–1200.
Lehmer, E.-M., & Bäuml, K.-H. T. (2018b). The many faces of part-list cuing - evidence for the interplay between detrimental and beneficial mechanisms. *Frontiers in Psychology, 9*, 701.
Nickerson, R. S. (1984). Retrieval inhibition from part-set cuing: A persisting enigma in

memory *research*. *Memory and Cognition, 12*, 531–552.
Raven, J. C. (1998). *Raven's progressive matrices and vocabulary scales.* Oxford: Pyschologists Press.
Roediger, H. L. I. I. I., Stellon, C. C., & Tulving, E. (1977). Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 174–188.
Roediger, H. L., III (1973). Inhibition in recall from cuing with recall targets. *Journal of Verbal Learning and Verbal Behavior, 36*, 644–657.
Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.
Rogers, B. (2001). *TOEFL CBT success.* Princeton, NJ: Peterson's.
Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior, 12*, 43–50.
Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology, 76*, 504–513.
Slamecka, N. J. (1969). Testing for associative storage in multitrial free recall. *Journal of Experimental Psychology, 81*, 557–560.
Sloman, S. A. (1991). Part-set cuing inhibition in category-instance and reason generation. *Bulletin of Psychonomic Society, 29*, 136–138.
Watkins, M. J. (1975). Inhibition in recall with extralist cues. *Journal of Verbal Learning and Verbal Behavior, 14*, 294–303.