



Deactivation of prospective memory intentions: Examining the role of the stimulus–response link

Emily Streeper¹ · Julie M. Bugg¹

Published online: 29 September 2020
© The Psychonomic Society, Inc. 2020

Abstract

Successful prospective remembering involves formation of a stimulus (e.g., bottle of medication and/or place where the bottle is kept)–response (e.g., taking a medication) link. We investigated the role of this link in the deactivation of no-longer-relevant prospective memory intentions, as evidenced by commission error risk. Experiment 1a contrasted two hypotheses of intention deactivation (*degree of fulfillment* and *response frequency*) by holding constant the degree of intention fulfillment (e.g., participants responded to one of two target words) while manipulating the number of times the intention was performed. Findings supported the response frequency hypothesis. Experiment 1b employed novel lure trials to examine what “stimulus” participants link the prospective memory response to—target words and/or the salient contextual cue—and compared commission errors to Experiment 1a. Findings suggested the salient context alone does not always function as the stimulus. Collectively these findings, in conjunction with those of Experiment 2 (a within-experiment replication) and a combined analysis, suggest that (a) intention deactivation is facilitated by prior responding (formation/strengthening of stimulus–response links), but additional research is needed to establish the robustness of this effect, and (b) when responding frequently to targets, participants are more likely to bind the response to the context alone than to the target or target/context combination, possibly because they learn to rely on context to predict target occurrence. The latter finding was robust and indicates that deactivation of the appropriate stimulus (target and/or context)–response link may be a critical component of reducing commission errors.

Keywords Prospective memory · Commission errors · Intention deactivation · Stimulus-response · Episodic traces

Prospective memory (PM) refers to the act of remembering to perform an intention in the future. In the past several decades, most PM research has addressed the question of how to successfully fulfill PM intentions and tried to understand what causes PM omission errors (i.e., failures to remember to perform an intention). In recent years, PM researchers have become increasingly interested in a different type of PM error: PM commission errors. A PM commission error is the act of erroneously repeating a PM intention when it is no longer relevant (e.g., erroneously taking medication that is no longer appropriate to take). Examining the underlying mechanisms

that cause commission errors is crucial to understanding why they happen and how to prevent them from occurring (see Möschl et al., 2020, for a recent review).

To examine PM commission errors in lab settings, the “finished paradigm” was developed (Scullin, Bugg, & McDaniel, 2012; for review, see Bugg & Streeper, 2019; cf. Walser, Fischer, & Goschke, 2012; for a related but distinct habitual PM paradigm, see Einstein, McDaniel, Smith, & Shaw, 1998; McDaniel, Bugg, Ramuschkat, Kliegel, & Einstein, 2009). In this paradigm, participants encounter two phases. In the first phase, referred to as the active PM phase, participants perform an ongoing task requiring word and nonword judgments (i.e., a lexical decision task). Along with completing the ongoing task, participants are given a PM intention to press a special key (e.g., the *Q* key) if they encounter either of two target words (e.g., *corn* or *dancer*). Participants are additionally instructed that target words will always appear on a salient, colored background (e.g., red screen). Most studies have used the four-target version of this paradigm in which each target word appears twice. Consequently, participants can fulfill the intention (respond to both target words at least once) in the

Portions of the data reported herein were presented at the 59th Annual Meeting of the Psychonomic Society in 2018

✉ Emily Streeper
streeper@wustl.edu

¹ Department of Psychological and Brain Sciences, Washington University in St. Louis, Campus Box 1125, St. Louis, MO 63130, USA

active PM phase. Following the active PM phase, participants are instructed that they no longer need to perform the special action in response to the target words—they simply should continue performing the ongoing task. These finished PM instructions are followed by the second phase of the paradigm, the finished PM phase. During this phase, participants again encounter the target words on the same salient background as during the active PM phase, but the target words are now irrelevant. Pressing the *Q* key in response to the no-longer-relevant target words indicates a commission error and suggests the PM intention is still accessible.

Most studies have examined commission errors following intention fulfillment (i.e., participants are presented with and respond to the target words in the active PM phase). However, a few studies using the finished paradigm have examined whether participants are inclined to perform a previously relevant intention they never had the opportunity to fulfill by employing a zero-target condition (Bugg & Scullin, 2013; Bugg, Scullin, & Rauvola, 2016; cf. Marsh, Hicks, & Bink, 1998). As in the four-target condition, participants encode the PM intention to press the *Q* key in response to the target words; however, in the zero-target condition, target words are never presented during the active PM phase. This means participants cannot fulfill the intention. Although it may seem intuitive that a PM intention performed multiple times (as in the four-target condition) would become somewhat habitual and therefore be harder to deactivate than an intention that was never performed (as in the zero-target condition), the findings from these studies were quite the opposite (see also Schaper & Grundgeiger, 2017). For example, Bugg and Scullin (2013) found that participants in the four-target condition deactivated the intention and did not make a commission error (but see Pink & Dodson, 2013). In striking contrast, 56% of participants in Experiment 1 and 46% of participants in Experiment 2 made a commission error in the zero-target condition. The authors concluded that PM intentions that remain unfulfilled are more accessible than intentions that are fulfilled, which is referred to as the intention fulfillment effect (Bugg & Streeper, 2019).

An important yet unanswered question concerns the cause(s) of the intention fulfillment effect. A few accounts have been proposed (Bugg & Scullin, 2013). One is the Zeigarnik (1938) account and refers to the possibility that the heightened accessibility reflects a Zeigarnik-like effect, whereby selectively in the zero-target condition participants experience tension about not fulfilling the intention and persevere on it. A second account is the episodic trace, or stop tag, account. According to this account, the act of pressing *Q* in response to target words during the active PM phase yields episodic traces of prior responding (cf. Hommel, 1998) and accordingly, a richer representation of intention completion. This enables participants to attach a “stop tag” to the intention when the finished instructions are shown, thereby facilitating creation of a no-go memory (cf. Hommel, Musseler,

Aschersleben, & Prinz, 2001; see Anderson & Einstein, 2017, for recent evidence for the stop-tag account). Because responding in the active PM phase occurs only in the four-target condition, this condition benefits from the stop tag while the zero-target condition does not.

Although it is difficult to disentangle these accounts when comparing the four-target and zero-target conditions, findings from one prior study that investigated the intermediate case of a partially fulfilled intention are informative. In their third experiment, Bugg and Scullin (2013) once again had participants encode the PM intention to press the *Q* key in response to the target words. However, in the active PM phase, only one of the two target words (e.g., *corn*) was shown, and it was presented once. (Hereafter, we refer to this as the one-target condition.) In the finished PM phase, they manipulated whether the first (now, no-longer-relevant) target word shown on the salient background was the presented word (i.e., *corn*) or the nonpresented word (i.e., *dancer*). The key finding was that commission errors were 3.5 times more likely when the nonpresented word was shown first in the finished PM phase, although this difference was marginal. Importantly, for present purposes, this finding provided preliminary support in favor of the episodic trace account as opposed to the Zeigarnik account. According to the Zeigarnik account, commission errors should have been equivalent for all participants in the one-target condition because all participants had the opportunity to fulfill the intention to the same degree (once) in the active PM phase. However, errors were lower for those participants that received the presented word first compared with those that received the nonpresented word first. The episodic trace account readily explains this difference. According to this account, prior responding yielded an episodic trace corresponding to the stimulus–response link of pressing *Q* in response to the presented word (*corn*) during the active PM phase, and thus a stop tag could be attached to this trace, facilitating deactivation. In contrast, there was no episodic trace corresponding to the stimulus–response link (i.e., *dancer*–press *Q*) for the nonpresented word.

Current study

To take stock, prior studies have suggested that the level of persisting activation of a no-longer-relevant intention is related to the degree of intention fulfillment. This is reflected in the intention fulfillment effect, as well as cross-experimental comparisons that have additionally considered commission error rates for partially fulfilled intentions (one-target condition), which appear to fall between the four-target and zero-target conditions (Bugg & Scullin, 2013; see Anderson & Einstein, 2017, for different results in a one-target condition in a paradigm where participants knew the task was completed after performing the action once). Additional evidence is needed to

inform theoretical accounts of the intention fulfillment effect and understand what factors affect the level of persisting activation of an intention once it is no longer relevant. Along these lines, the current study aimed to take a closer look at the stimulus–response link—that is, how the intention is represented. This link is purported to play a central role in the episodic trace account of commission errors.

Experiment 1a examined whether differences in commission error risk are due to the degree of intention fulfillment per se or the total number of responses that were made to a target stimulus in the active PM phase (i.e., strengthening of the stimulus–response link). Indeed, these two factors have covaried in prior experiments contrasting four-target (fulfilled intention/4 responses), one-target (partially unfulfilled intention/1 response), and zero-target (unfulfilled intention/0 responses) conditions.¹ Consequently, it has not been possible to tease apart their effects. Experiment 1b examined what comprises the stimulus component of the stimulus–response link using the novel approach of embedding lure trials in the finished PM phase, and comparing commission errors between Experiment 1b (lure trials) and Experiment 1a (standard trials). Addressing this question is important for further informing the episodic retrieval account and understanding the conditions that may increase susceptibility to commission errors. Experiment 2 further contrasted these conditions, albeit head-to-head within a single experiment, thereby offering an opportunity to replicate the patterns observed across Experiments 1a and 1b.

Experiment 1a

The focus of Experiment 1a was to contrast two hypotheses that fall out of two extant accounts of the intention fulfillment effect: the Zeigarnik account and the episodic trace account. One hypothesis, termed here the *degree of fulfillment hypothesis*, posits that commission errors should be least likely when an intention is fulfilled meaning a participant has responded to all targets at least once, more likely when an intention is partially fulfilled, meaning a participant has responded to a subset of the targets at least once, and most likely when the intention is unfulfilled meaning a participant has responded to no targets. This hypothesis falls out of the Zeigarnik account (Bugg & Scullin, 2013) in that the degree of intention accessibility in the finished PM phase is predicted to be higher to the extent that intentions are left unfulfilled, as this may lead to perseverating on the unfulfilled intentions. The second hypothesis, termed here the *response frequency hypothesis*, posits that commission errors should be less likely

the more frequently a participant performs the intention in the active PM phase. This hypothesis falls out of the episodic trace account (Bugg & Scullin, 2013) in that responding more frequently to a target word should create a stronger stimulus–response link (solidify intention representation through the accumulation of traces) and make it easier to associate a stop tag with this representation, leading to better intention deactivation.

To contrast these hypotheses, we compared performance in a one-target condition (e.g., participants encoded an intention to respond to both *corn* and *dancer*, but performed the PM intention only once in the active PM phase for *corn*), which necessarily represented partial completion of the intention, to a novel four single-target condition that also represented partial completion of the intention (e.g., participants encoded an intention to respond to both *corn* and *dancer*, but performed the PM intention four times in the active PM phase for *corn* only). Critically, comparing these two conditions enabled us to hold constant the degree of intention fulfillment (i.e., only one of the two target words was responded to at least once in both conditions) while varying the number of times participants responded to a target word (one vs. four, respectively). We also included a zero-target condition as a theoretically interesting comparison. This allowed us to examine the effects of partial intention fulfillment (one-target and four single-target conditions) relative to a completely unfulfilled intention. The experimental procedure for these three conditions is depicted in Fig. 1.

Theoretically, the critical comparison of interest is between the one-target and four single-target conditions. According to the degree of fulfillment hypothesis, commission error rates should be comparable between these two conditions because both involve partial fulfillment. In contrast, according to the response frequency hypothesis, the four single-target condition should have significantly lower commission error rates compared with the one-target condition. Regarding the other potential comparisons, the degree of fulfillment hypothesis posits that commission error rates should be higher in the zero-target condition, which represents an unfulfilled intention, compared with the one-target condition and four single-target condition. The same prediction holds for the response frequency hypothesis, although it would attribute the difference to the number of times participants responded in each condition. As in our prior research (e.g., Bugg & Scullin, 2013; Bugg et al., 2016) the primary dependent variable was the number of participants in each condition that made a commission error during the finished PM phase to gauge accessibility of the encoded intention (which was to press *Q* to either *corn* or *dancer*). Additionally, we examined the number of participants who made an error selectively on the first target, which was always the target that was previously responded to in the one-target and four-target conditions (e.g., *corn* in examples above).

¹ In Bugg and Scullin (2013), the number of responses on average in the four-target condition in the active PM phase was 3.9 out of 4. The number of correct PM responses in the one-target condition was .93 out of 1.

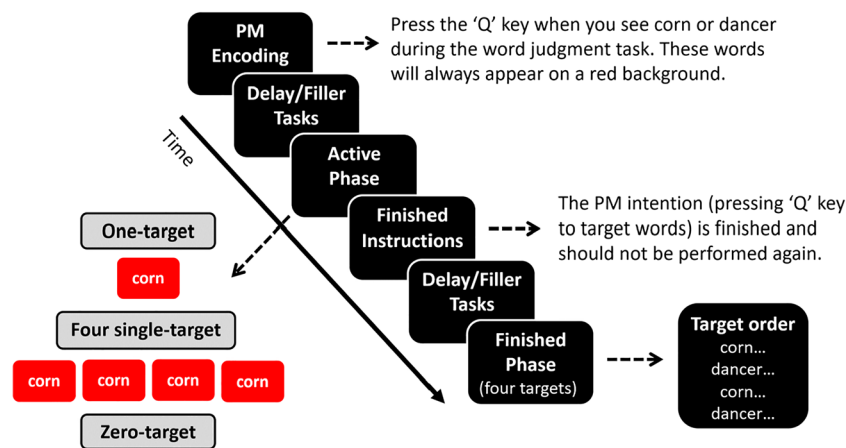


Fig. 1 Experimental procedure for Experiment 1a. In the active PM phase, participants encountered one encoded target word once (one-target condition), one encoded target word four times (four single-target condition), or neither encoded target word (zero-target condition)

Method

Design and participants Seventy-seven Washington University in Saint Louis undergraduate students, with normal or corrected-to-normal vision and color vision, and who reported English as their native language, participated in this study for either monetary compensation or course credit. The one-target and four single-target conditions were run simultaneously, and participants were randomly assigned to one of these conditions. The zero-target condition was run as a dangling comparison condition. A priori, we implemented a stopping rule of 24 participants per condition (following Bugg & Scullin, 2013) who met PM performance-based inclusion requirements. Those requirements were that participants in the zero-target condition should not have pressed *Q* in the active PM phase and participants in the four single-target condition should have pressed *Q* more than once to the single target presented four times in the active PM phase. We reasoned that a participant that responded twice or three times in the four single-target condition still met the purpose of the condition (partial fulfillment) and responded more frequently than a participant in the one-target condition and thus should be included.² Later, we similarly decided to exclude participants in the one-target condition who did not press the *Q* key in response to the single target in the active PM phase. We reasoned that a participant who did not respond once in the one-target condition (i.e., responded zero times) no longer met the purpose of this condition (partial fulfillment) and instead more closely mimicked the zero-target condition (although a target was never shown in the active PM phase of that condition). In the one-target condition, two participants were excluded for

failing to press the *Q* key in response to the target word in the active PM phase.

In addition, three participants were excluded for either failing to read instructions (i.e., one participant in the four single-target condition pressed through the finished instructions without reading them), or understand instructions (i.e., one participant in the one-target condition and one in the four single-target condition did not know how to advance to the next trial when the target appeared in the finished PM phase). The final sample for Experiment 1a ($N = 72$, 24 per condition) was 73.6% female (one participant did not report sex).

Materials and procedure The procedure is shown in Fig. 1. First, all participants were instructed to use only one hand when responding during the task. Then, they were given the opportunity to practice the ongoing lexical decision task for eight trials. They indicated whether they thought the letters onscreen were a word or a nonword by pressing the labeled *Y* (5) or *N* (6) keys on the number pad. After practice, participants were given the PM intention instructions. They were instructed to press the *Q* key if they saw either of the two target words, which would appear on the colored background. There were two possible sets of target words (*corn/dancer* and *fish/writer*) and two possible background colors (red and blue). The target words and background colors were counterbalanced across participants. Participants were told they could press the *Q* key in response to target words before or after making their lexical decision. All participants regardless of condition were given two target words to encode.

Once participants finished reading these instructions, they wrote down their target words and completed a demographic form and vocabulary test to create an ~5-min delay between encoding and testing (Einstein & McDaniel, 1990). After completing these forms, participants began the active PM phase. The active PM phase consisted of 76 trials. In the one-target condition, participants only saw one of their target

² In the four single-target condition, two participants pressed the *Q* key in the active PM phase three times. All other participants in the four single-target condition responded four times in the active PM phase.

words one time (e.g., only *corn* one time) on Trial 38. In the four single-target condition, participants also saw only one of their target words, but they saw that target word four times (e.g., only *corn* four times) on Trials 14, 33, 52, and 71. For the zero-target condition, participants did not encounter either target word in the active PM phase.

Upon completing the active PM phase, all participants were given the following finished instructions before beginning the finished PM phase: “PLEASE NOTE THAT YOU NO LONGER NEED TO PRESS ‘Q’ IN THE PRESENCE OF TARGET WORDS. THAT TASK IS FINISHED AND SHOULD NOT BE PERFORMED AGAIN. Just as before, you will determine whether a string of letters forms a word or a nonword by pressing the keys marked Y and N on the number pad. YOUR ONLY GOAL is to make word/nonword judgments.” In the finished PM phase, all participants experienced a brief delay, during which they first completed a short block of lexical decision trials (24 trials with no targets; see e.g., Bugg & Scullin, 2013; Bugg and Scullin, 2013; Bugg et al., 2016; Scullin et al., 2012), and then another vocabulary form (which differed from the first one). Following this ~5-min delay, participants completed a 118-trial lexical decision block that included four trials in which they encountered the no-longer-relevant target words (e.g., *corn* and *dancer* each presented twice). In the one-target condition and the four single-target condition, the first no-longer-relevant target word presented in the finished PM phase was the word they saw in the active PM phase (e.g., *corn*). In the zero-target condition, the finished PM phase matched that of the one-target condition and the four single-target condition. In conditions where the target words were *corn* and *dancer*, the target words appeared in the 118-trial lexical decision block on Trials 42, 66, 90, and 113. In conditions where the target words were *fish* and *writer*, the target words appeared on Trials 39, 47, 83, and 103. After completing the finished PM phase, participants completed a postexperimental questionnaire.

Results

Active PM phase PM hits Following Bugg et al. (2016), PM hits were defined as a Q press that occurred on the target trial or within two trials after the presentation of the target.³ An independent *t* test showed a significant difference in average number of PM hits between the one-target condition ($M = 1.00$, $SD = .00$) and the four single-target condition ($M = 3.92$, $SD = .28$), $t(23) = 50.61$, $p < .001$, as was expected given

³ Following past research (Bugg et al., 2016; Scullin et al., 2012), we examined when PM hits occurred in relation to the target words. Across conditions, 88.14% of PM hits occurred on the presentation of the target word, while 11.02% occurred on the subsequent trial and .85% occurred on the trial after the subsequent trial. There were no false alarms.

the difference in target presentation between conditions and the inclusion criteria.⁴

Finished PM phase commission errors Following prior research (Bugg et al. 2016) a commission error was defined as a Q press that occurred in the finished PM phase,⁵ and our primary interest was the effect of condition on the number of participants who made at least one commission error (see Fig. 2). Significantly fewer participants made a commission error in the four single-target condition (21%) compared with the one-target condition (50%), $\chi^2(1) = 4.46$, $p = .035$. The number of participants who made a commission error in the zero-target condition (75%) was not significantly higher than the one-target condition, $\chi^2(1) = 3.20$, $p = .074$, but was significantly higher than the four single-target condition, $\chi^2(1) = 14.11$, $p < .001$.

Following Bugg et al. (2016), we also examined the number of participants who made a commission error on the first no-longer-relevant target presented in the finished PM phase. In the one-target and four single-target conditions, the first no-longer-relevant target was the same target participants responded to in the active PM phase. The patterns mirrored those found when considering commission errors to any target (preceding analysis) though in this analysis the difference between the four single-target condition (21%) and the one-target condition (46%) was not significant, $\chi^2(1) = 3.38$, $p = .066$. The number of participants who made a commission error on the first no-longer-relevant target in the zero-target condition (71%) was not significantly higher than the one-target condition, $\chi^2(1) = 3.09$, $p = .079$, but was significantly higher than the four single-target condition, $\chi^2(1) = 12.08$, $p = .001$.

Lexical decision task performance To examine the difference in speeding from the active PM phase, a phase that includes a PM task, to the finished PM phase, a phase that consists only of the ongoing lexical decision task, by condition, we examined reaction times (RTs) in milliseconds (ms) on nontarget trials in the ongoing lexical decision task in both phases (see Table 1). We restricted our analyses to correct trials and trials that did not occur within three trials after a target was presented, and only RTs within 2.5 standard deviations from each participant’s mean for each phase were included in analyses (cf. Bugg & Ball, 2017; Lourenço, White, & Maylor, 2013). A 2 (phase: active PM phase, finished PM phase) \times 3 (condition: one-target, four single-target, zero-target) mixed-model ANOVA showed a significant effect of phase, $F(1, 69) =$

⁴ A Levene’s test indicated unequal variances ($F = 10.12$, $p = .003$). Degrees of freedom were adjusted from 46 to 23.

⁵ Following past research (Bugg et al., 2016; Scullin et al., 2012), we examined when PM commission errors occurred in relation to the target words. Across conditions, 90.99% of commission errors occurred on the presentation of the no-longer-relevant target word, while 6.31% occurred on the subsequent trial and no commission errors occurred on the trial after the subsequent trial. Additionally, 2.70% of commission errors occurred more than two trials after the target words.

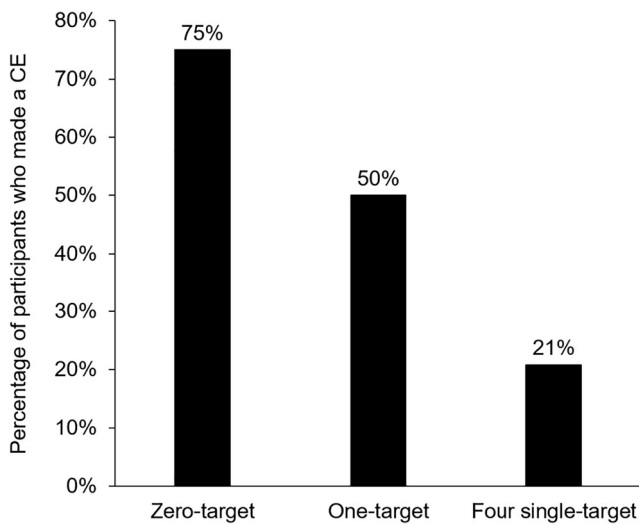


Fig. 2 Percentage of participants who made at least one commission error by condition in Experiment 1a

23.99, $p < .001$, but not an effect of condition, $F(2, 69) = .64$, $p = .528$, or a significant Phase \times Condition interaction, $F(2, 69) = 1.68$, $p = .193$. These results suggest participants in all conditions sped up from the active PM phase ($M = 716$, $SD = 125$) to the finished PM phase ($M = 575$, $SD = 266$) at similar rates.

As for average accuracy, a 2 (phase: active PM phase, finished PM phase) \times 3 (condition: one-target, four single-target, zero-target) mixed-model ANOVA showed significant effects of phase, $F(1, 69) = 143.78$, $p < .001$, and condition, $F(2, 69) = 4.48$, $p = .015$, but not a significant Phase \times Condition interaction, $F(2, 69) = 2.39$, $p = .099$ (see

Table 1 Mean RTs in milliseconds on the ongoing lexical decision task in active PM and finished PM phase by condition

| Experiment | Condition | Active PM phase M (SD) | Finished PM phase M (SD) |
|------------|-------------------------|------------------------------|--------------------------------|
| 1a | Zero-target | 766 (161) | 583 (174) |
| | One-target | 676 (100) | 609 (427) |
| | Four single-target | 707 (90) | 532 (55) |
| 1b | Lure zero-target | 728 (111) | 560 (59) |
| | Lure one-target | 663 (94) | 516 (45) |
| | Lure four single-target | 714 (96) | 512 (55) |
| 2 | Modified zero-target | 679 (99) | 552 (80) |
| | One-target | 696 (119) | 534 (111) |
| | Four single-target | 709 (96) | 558 (92) |
| | Lure zero-target | 715 (166) | 548 (64) |
| | Lure one-target | 675 (88) | 542 (65) |
| | Lure four single-target | 717 (100) | 558 (48) |

Note. Reaction time analyses were conducted on correct trials and were trimmed 2.5 standard deviations from each participant’s mean for each block (per Bugg & Ball, 2017; Lourenço et al., 2013).

Table 2).⁶ Participants performed worse on the ongoing lexical decision task in the active PM phase ($M = 82.91\%$, $SD = 7.58\%$) than in the finished PM phase ($M = 91.63\%$, $SD = 5.66\%$). Post hoc Tukey HSD tests indicated participants performed significantly better in the four single-target condition ($M = 89.24\%$, $SD = 3.47\%$), compared with the one-target condition ($M = 84.55\%$, $SD = 7.53\%$), $p = .014$. Average accuracy did not differ significantly between the one-target condition and the zero-target condition ($M = 88.02\%$, $SD = 5.14\%$), $p = .090$, or between the four-target condition and the zero-target condition, $p = .738$.

Discussion

The key novel finding from Experiment 1a was that fewer participants made at least one commission error in the four single-target condition compared with the one-target condition. That is, in conditions that were equated in the degree of fulfillment (i.e., participants in both conditions responded only to one of the two target words), the condition in which responding occurred more frequently led to fewer participants making at least one commission error. This pattern supports the response frequency hypothesis over the degree of fulfillment hypothesis and highlights the role of responding in intention deactivation. It appears that the more frequently one fulfills an intention, the less accessible the intention will be later, and this is true even under conditions that control for the degree of fulfillment. This accords with the view that episodic traces of prior responding facilitate the linking of a stop tag to a no-longer-relevant intention. Presumably, the strength of these traces is enhanced with repeated responding such that the stronger the stimulus–response link (the more robust the representation of the intention), the higher the likelihood of effectively binding a stop tag to the intention, and thereby deactivating the intention.

Surprisingly, and contrary to both accounts, commission error rates did not significantly differ between the one-target condition and the zero-target condition. However, the direction of the difference was as predicted (by both the response frequency and degree of fulfillment accounts) with higher commission error rates in the zero-target condition compared with the one-target condition. Possibly this could reflect inadequate power, and we will return to this possibility following Experiment 2.

Finally, Experiment 1a produced an intention fulfillment effect in the form of higher rates of commission errors in the zero-target condition compared with the four single-target condition, consistent with both accounts. In prior studies demonstrating the intention fulfillment effect, the four-target condition comprised presentation of both targets (twice each) and thus complete fulfillment of the intention (Bugg & Scullin,

⁶ A Levene’s test indicated unequal variances for accuracy in the finished PM phase ($F = 4.42$, $p = .016$).

Table 2 Mean accuracy on the ongoing lexical decision task in active PM and finished PM phase by condition

| Experiment | Condition | Active PM phase <i>M%</i> (<i>SD%</i>) | Finished PM phase <i>M%</i> (<i>SD%</i>) |
|------------|-------------------------|------------------------------------------|--------------------------------------------|
| 1a | Zero-target | 84.48 (6.96) | 91.57 (4.80) |
| | One-target | 79.11 (8.39) | 89.99 (7.86) |
| | Four single-target | 85.13 (5.96) | 93.34 (2.88) |
| 1b | Lure zero-target | 84.87 (5.29) | 93.19 (3.39) |
| | Lure one-target | 84.17 (7.70) | 91.11 (5.18) |
| | Lure four single-target | 85.53 (6.24) | 92.23 (3.17) |
| 2 | Modified zero-target | 84.76 (7.71) | 93.45 (3.63) |
| | One-target | 82.96 (8.82) | 90.83 (8.75) |
| | Four single-target | 83.45 (6.43) | 91.16 (6.12) |
| | Lure zero-target | 84.48 (6.56) | 94.00 (3.07) |
| | Lure one-target | 83.94 (8.80) | 92.84 (5.43) |
| | Lure four single-target | 85.52 (5.44) | 92.97 (3.63) |

2013; Bugg et al., 2016). The current finding extends this effect to a partially completed intention that was responded to repeatedly (i.e., the four single-target condition).

Collectively, the findings of Experiment 1a point to a continuum of intention deactivation that corresponds to the degree of intention fulfillment, with initial evidence suggesting that in intermediate conditions (conditions of partial fulfillment), intention deactivation is greater the more frequently a PM target has been responded to previously (i.e., in the four single-target condition as compared with the one-target condition). We interpret this result to suggest an important role for the strengthening of the stimulus–response link in intention deactivation, consistent with the episodic trace account.

Experiment 1b

While Experiment 1a identified the strength of the stimulus–response link as a factor that affects intention deactivation independently of the degree of intention fulfillment by manipulating the number of responses made to a single target word, Experiment 1b aimed to examine what precisely constitutes the “stimulus” portion of the stimulus–response link and what the implications are for intention deactivation. As previously mentioned, in the finished paradigm participants are informed that target words will appear in a salient context (i.e., uniquely colored background) when they encode the PM intention. Later, when the no-longer-relevant targets are presented in the finished PM phase, that same salient context reappears and the representation of this context appears to be important for eliciting commission errors (Scullin et al., 2012; Scullin, Bugg, McDaniel, & Einstein, 2011). It has been suggested that the

salient context may cause the PM intention to “pop” into mind in the finished PM phase (i.e., to be spontaneously retrieved; Scullin & Bugg, 2013). In other words, during the active PM phase participants may be associating the PM response not necessarily with the target word but potentially with the salient context, given that it accompanies the target 100% of the time (i.e., it is 100% predictive of target occurrence). However, this is merely speculation as no study has directly examined whether participants are linking the PM response to the target word itself, some combination of the target word and salient context *or* if they might indeed be relying primarily on the salient context to guide intention retrieval.

This question of how the PM intention is represented, and in particular what constitutes the stimulus in the stimulus–response link, has important implications for understanding and predicting when commission errors will occur. If participants are primarily relying on the salient context, the implications are that (a) commission errors can occur even when the target stimulus (word) is absent so long as participants are in the salient context associated with the intention (i.e., in the real-world, if one’s bathroom is the context, participants may make a commission error by taking medication B upon walking into the bathroom even if medication A is no longer there), and (b) intention deactivation that focuses on only the target itself will not be sufficient to prevent commission errors. Theoretically speaking, considering the episodic trace account, this means that applying a stop tag to the association between a specific target word and a response (i.e., in the real world, to the association between medication A and the response), may not be sufficient to prevent one from making a commission error when the word is re-presented on the salient background in the finished PM phase (i.e., when medication A is again encountered in the bathroom).

To examine this question, we modified the paradigm used in Experiment 1a to include a “lure word” in the finished PM phase (see Fig. 3). Lure words were *not* the target words encoded in the

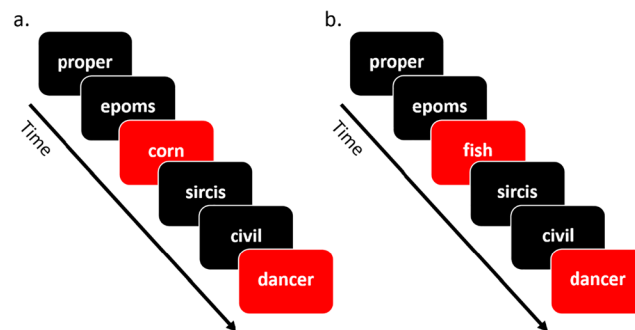


Fig. 3 Finished PM phase in Experiments 1a and 1b. The left panel (a) depicts the finished PM phase from Experiment 1a. Note that *corn*, the first word presented on a red screen, was in fact a target word. The right panel (b) depicts the lure condition where the first “target” trial presents a lure (i.e., nontarget) word instead of one of the actual target words. In both experiments, these words appeared on the encoded (i.e., correct) background color, in this case, red. (Color figure online)

active PM phase, but they nonetheless appeared on the same colored background previously linked only to target words. For example, if the target words (e.g., *corn*) were shown on a red background in the active PM phase, then the lure word (e.g., *fish*) would also be shown on a red background. Examining how participants respond to the lure words informs the question of what plays the role of the stimulus in the stimulus–response link. If participants are linking their PM response to the salient context, regardless of condition (four single-target, one-target, or zero-target), they should show similar commission error rates for the lure words as they do for the previously relevant target words because the identity of the word should not impact whether they respond (press *Q*), only the background color should matter. However, if they are linking the PM response to the target word alone or a combination of the salient background and the target word (i.e., if the target word itself plays some role in the stimulus component of the stimulus–response link), then regardless of condition the commission error rates for lure words should be lower than the rates for the previously relevant target words. To test these hypotheses, we compared commission error rates between the lure conditions from Experiment 1b and their control (nonlure) counterparts (with an actual, encoded target word) from Experiment 1a. In this experiment, the dependent variable of interest was selectively the commission error rates to the first “target” in the finished PM phase because the lure word occurred only on the first target trial in the lure conditions, and thus was compared with the nonlure target word that was presented on the first trial in the control conditions.

Method

Design and participants Seventy-five Washington University in Saint Louis undergraduate students with normal or corrected-to-normal vision and color vision, and who reported English as their native language, participated in this study for either monetary compensation or course credit. Participants were randomly assigned to the lure one-target condition, the lure four single-target condition, and the lure zero-target condition (or the zero-target condition from Experiment 1a). The stopping rule and exclusion criteria from Experiment 1a (i.e., 24 participants who met inclusion criteria per condition) were applied in Experiment 1b. In the lure one-target condition, two participants were excluded for failing to press the *Q* key to the single target in the active PM phase.⁷ In addition, one participant was excluded for repeatedly falling asleep in the lure four single-target condition. The final sample for Experiment 1b ($N = 72$, 24 per condition) was 63.9% female.

⁷ In the lure four single-target condition, four participants pressed the *Q* key in the active PM phase three times and two participants pressed the *Q* key in the active PM phase two times. All other participants in the lure four single-target condition responded four times in the active PM phase.

Materials and procedure The same materials and procedure for the conditions in Experiment 1a were used for the conditions in Experiment 1b (see Fig. 1), with one modification. For all the conditions in Experiment 1b (i.e., lure one-target condition, lure four single-target condition, and lure zero-target condition), participants were presented with one lure word (e.g., *fish* when the target words encoded were *corn* and *dancer*) during the finished PM phase (see Fig. 3). The lures appeared on the same colored background on which participants were told the target words would appear (e.g., red; see Fig. 3). For the lure one-target and lure four single-target conditions, the lure was later followed by one presentation of the target word previously shown in the active PM phase (e.g., *corn*) and two presentations of the other target word (e.g., *dancer*). For the lure zero-target condition, the finished PM phase after the lure word matched that of the lure one-target and lure four single-target conditions.

Results

For each of the analyses below, we applied the same criteria used in Experiment 1a for determining PM hits and commission errors, as well as in the analysis of RT and accuracy.

Active PM phase PM hits⁸ The average number of PM hits was equivalent between the control one-target condition and the lure one-target condition ($M = 1.00$, $SD = .00$), as expected given exclusionary criteria. An independent *t* test indicated no significant difference in the average number of PM hits between the control four single-target condition ($M = 3.92$, $SD = .28$) and the lure four single-target condition ($M = 3.67$, $SD = .64$), $t(31.70) = 1.76$, $p = .088$.⁹

Finished PM phase commission errors¹⁰ Unlike in Experiment 1a, the single commission error measure of interest was the number of participants who made a commission error on the first “target” shown in the finished PM phase (see Fig. 4). This was because the lure (in the lure conditions) appeared only on the first trial and not later trials. We compared the lure conditions (whose first “target” was a lure trial) from Experiment 1b to the control conditions (whose first “target” was in fact a target) from Experiment 1a. For the four single-target condi-

⁸ Across lure conditions, 78.57% of PM hits occurred on the presentation of the target word, while 21.43% occurred on the subsequent trial. No PM hits occurred after the subsequent trial. There were no false alarms.

⁹ A Levene’s test indicated unequal variances ($F = 14.45$, $p < .001$). Degrees of freedom were adjusted from 46 to 31.70.

¹⁰ Across lure conditions, 75.00% of commission errors occurred on the presentation of the first “target” word in the finished PM phase, while 16.67% occurred on the subsequent trial and no commission errors occurred on the trial after the subsequent trial. Additionally, 8.33% occurred more than two trials after the first “target” word in the finished PM phase.

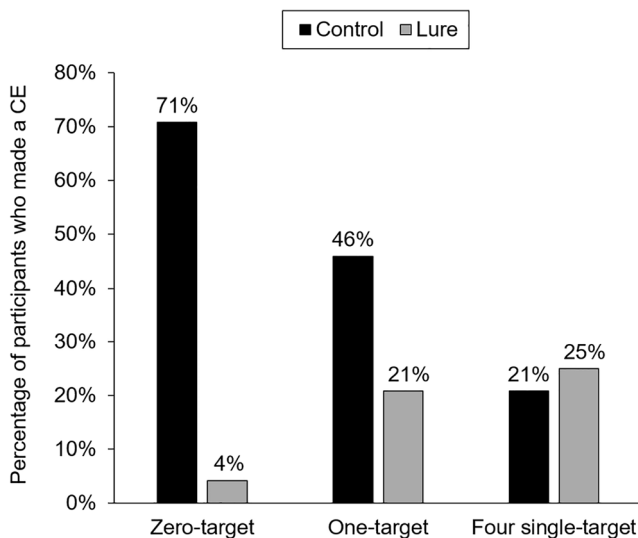


Fig. 4 Percentage of participants who made a commission error (CE) on the first “target” presented in the finished PM phase by condition. The control conditions are from Experiment 1a and the first “target” was in fact a target on a salient background; the lure conditions are from Experiment 1b, and the first “target” was a lure trial (nontarget word) on a salient background

tions, the number of participants who made a commission error did not differ between the control (21%) and lure condition (25%), $\chi^2(1) = 0.12$, $p = .731$. For the one-target conditions, numerically more participants made a commission error on the first target in the control one-target condition (46%) compared with its lure counterpart (21%); however, this difference was not significant, $\chi^2(1) = 3.38$, $p = .066$. In the zero-target condition, significantly more participants made a commission error in the control condition (71%) compared with the lure condition (4%), $\chi^2(1) = 22.76$, $p < .001$. Strikingly, only one participant made a commission error on the first “target” in the lure zero-target condition compared with 17 participants in the control zero-target condition.

Lexical decision task performance We examined RTs during the active PM phase and the finished PM phase by performing a 2 (phase: active PM phase, finished PM phase) \times 3 (number of targets: zero, one, four) \times 2 (first target type condition: lure, control) mixed-model ANOVA (see Table 1). There was a significant effect of phase, $F(1, 138) = 109.07$, $p < .001$, indicating speeding from the active PM phase ($M = 709$, $SD = 115$) to the finished PM phase ($M = 552$, $SD = 193$).¹¹ No other main effects or interactions were significant. These patterns suggest participants sped up from the active PM phase to the finished PM phase comparably across all conditions.

As for average accuracy, a 2 (phase: active PM phase, finished PM phase) \times 3 (number of targets: zero, one, four) \times 2 (first target type condition: lure, control) mixed-model

ANOVA showed significant effects of phase, $F(1, 138) = 287.00$, $p < .001$, and number of targets, $F(2, 138) = 4.42$, $p = .014$, but no significant effect of first target type condition, $F(1, 138) = 2.06$, $p = .154$ (see Table 2).¹² None of the interactions were significant. Participants performed worse on the ongoing lexical decision task in the active PM phase ($M = 83.88\%$, $SD = 7.06\%$) than in the finished PM phase ($M = 91.90\%$, $SD = 4.91\%$). Post hoc Tukey HSD tests indicated participants performed significantly better in the four single-target conditions ($M = 89.06\%$, $SD = 3.85\%$), compared with the one-target condition ($M = 86.09\%$, $SD = 6.79\%$), $p = .017$. However, average accuracy did not significantly differ between the one-target condition and the zero-target condition ($M = 88.53\%$, $SD = 4.62\%$), $p = .061$, or between the four-target condition and the zero-target condition, $p = .872$.

Discussion

The purpose of Experiment 1b was to determine what constitutes the stimulus portion of the stimulus–response link. For the zero-target condition, the commission error rate on the first “target” in the finished PM phase was lower in the lure condition compared with the control condition. While not significant, the commission error rate was also numerically lower in the lure condition compared with the control condition for the one-target condition. These findings suggest that participants in the zero-target condition linked their PM intention to either the target word itself or a combination of the target word and the salient context, a tendency that was apparent though not as robust for participants in the one-target condition. A lower rate of commission errors in a lure condition compared with the corresponding control condition implies that the stimulus portion of the stimulus–response link comprised more than just the salient context. If it was just the salient context, then commission error rates should have been equivalent for the lure and control conditions because the context was present in both cases.

Interestingly, in the four single-target condition, commission error rates on the first “target” in the finished PM phase were equivalent for the lure and control conditions. This suggests that in the four single-target condition, the stimulus participants linked their PM intentions to may have been the salient context alone. If the stimulus comprised the word in some form (on its own or in conjunction with the salient context), then commission error rates should have been higher in the control condition. The implication is that, in these cases, the salient context alone may be enough to trigger retrieval of the intention and tempt participants into committing commission errors. Before discussing these implications further, we first attempt to replicate these patterns in another experiment.

¹¹ A Levene’s test indicated unequal variances for the reaction times in the finished PM phase ($F = 2.59$, $p = .029$).

¹² A Levene’s test indicated unequal variances for accuracy in the finished PM phase ($F = 3.53$, $p = .005$).

Experiment 2

The contrast between the control (nonlure) conditions of Experiment 1a and lure conditions of Experiment 1b provided novel evidence demonstrating that, in some cases participants associate the PM response not with a specific target or a target/context conjunction, but merely with the salient context in which the intention is performed. The main purpose of Experiment 2 was to try to replicate the patterns observed in the cross-experimental contrast between Experiments 1a and 1b. Toward this end, in Experiment 2 we randomly assigned participants to one of the six conditions that comprised Experiments 1a and 1b. A related purpose was to collect additional data to allow for a higher-powered test of our primary hypotheses in an analysis that combined the data from all experiments.

A final purpose regarded the zero-target condition. In Experiments 1a and 1b all participants (regardless of condition) received the same instructions at the end of the finished PM phase to not perform the PM task “again.” As a reviewer noted, it may be possible that participants in the zero-target condition were especially inclined to press the *Q* key in the finished PM phase (as indicated by higher rates of commission errors) because the instructions to not perform the task “again” may have led them to think the experimenter made an error since they never actually responded previously (i.e., no targets were shown). To address this possibility, in Experiment 2, all conditions were identical to Experiments 1a and 1b, except the zero-target conditions, which we modified by eliminating the word *again* from the finished instructions.

Method

Design and participants One hundred and forty-two Washington University in Saint Louis undergraduate students with normal or corrected-to-normal vision and color vision, and who reported English as their native language, participated in this study for course credit. Participants were randomly assigned to the control one-target condition, the control four single-target condition, the control zero-target modified condition, the lure one-target condition, the lure four single-target condition, and the lure zero-target modified condition. The stopping rule and exclusion criteria from Experiment 1a and 1b were adopted in Experiment 2. In the control one-target condition, one participant was excluded for failing to press the *Q* key to the single target in the active PM phase.¹³ In addition, one participant in the lure four single-target condition

¹³ In the control four single-target condition, two participants pressed the *Q* key in the active PM phase three times, while all other participants responded four times in the active PM phase. In the lure four single-target condition, all participants pressed the *Q* key in the active PM phase four times.

was excluded for failing to complete the experiment. Due to the COVID-19 pandemic, we were unable to meet the 24-participant stopping goal for two conditions. The lure four single-target condition consists of data from 21 participants, and the control one-target condition consists of 23 participants. All other conditions consisted of 24 participants. The final sample for Experiment 2 ($N = 140$) was 76.4% female.

Materials and procedure The same materials and procedure for the conditions in Experiment 1a and Experiment 1b were used for the conditions in Experiment 2 (see Figs. 1 and 3), with one modification. For the control zero-target and lure zero-target modified conditions, the finished instructions were slightly modified to exclude the word “AGAIN” and read: “PLEASE NOTE THAT YOU NO LONGER NEED TO PRESS ‘Q’ IN THE PRESENCE OF TARGET WORDS. THAT TASK IS FINISHED AND SHOULD NOT BE PERFORMED. Just as before, you will determine whether a string of letters forms a word or a nonword by pressing the keys marked Y and N on the number pad. YOUR ONLY GOAL is to make word/nonword judgments.” No other changes were made to any of the conditions.

Results

For each of the analyses below, we applied the same criteria used in Experiments 1a and 1b for determining “hits” and commission errors, as well as in the analysis of RT and accuracy.

Active PM phase PM hits¹⁴The average number of PM hits was equivalent between the control one-target condition and the lure one-target condition ($M = 1.00$, $SD = .00$), as expected. Also, as expected, an independent *t* test showed a significant difference in average number of PM hits between the control one-target condition ($M = 1.00$, $SD = .00$) and the control four single-target condition ($M = 3.92$, $SD = .28$), $t(23) = 50.61$, $p < .001$.¹⁵ An independent *t* test indicated no significant difference in the average number of PM hits between the control four single-target condition ($M = 3.92$, $SD = .28$) and the lure four single-target condition ($M = 4.00$, $SD = .00$), $t(23) = 1.45$, $p = .162$.¹⁶

¹⁴ Across control conditions, 76.07% of PM hits occurred on the presentation of the target word, while 23.93% occurred on the subsequent trial. No PM hits occurred after the subsequent trial. Across lure conditions, 75.93% of PM hits occurred on the presentation of the target word, while 24.07% occurred on the subsequent trial. No PM hits occurred after the subsequent trial. There were no false alarms.

¹⁵ A Levene’s test indicated unequal variances ($F = 9.69$, $p = .003$). Degrees of freedom were adjusted from 45 to 23.

¹⁶ A Levene’s test indicated unequal variances ($F = 8.83$, $p = .005$). Degrees of freedom were adjusted from 43 to 23.

Finished PM phase commission errors¹⁷The same analyses employed in Experiments 1a and 1b were conducted in Experiment 2.

Control conditions (effects of response frequency manipulation). As in Experiment 1a, our primary variable of interest was the number of participants who made a commission error in the finished PM phase (see Fig. 5). The number of participants who made a commission error in the four single-target condition (42%) was nominally lower than the one-target condition (52%), $\chi^2(1) = .52$, $p = .471$, but this difference was not significant (unlike in Experiment 1a). Additionally, the number of participants who made a commission error in the zero-target modified condition (54%) was not significantly greater than the one-target condition, $\chi^2(1) = .02$, $p = .891$, or the four single-target condition, $\chi^2(1) = .75$, $p = .386$. As can be seen in Fig. 5, the percentage of participants who made at least one commission error in the zero-target modified condition was low compared with the standard zero-target condition in Experiment 1a,¹⁸ whereas the commission error rate in the one-target condition was comparable between experiments.

For the analysis that was restricted to the first target in the finished PM phase, there were no differences in the number of participants who made a commission error between any of the conditions: one-target condition (48%) versus four single-target condition (38%), $\chi^2(1) = .51$, $p = .474$; one-target condition versus zero-target modified condition (46%), $\chi^2(1) = .02$, $p = .891$; zero-target modified condition versus four single-target condition, $\chi^2(1) = .34$, $p = .558$.

Lure conditions (effects of lure manipulation). As in Experiment 1b, our primary variable of interest was the number of participants who made a commission error on the first “target” that was either a lure (lure conditions) or the no-longer-relevant target word (control). Mirroring the between-experiment contrast, the number of participants who made a commission error in the four single-target conditions did not differ between the control (38%) and lure condition (19%), $\chi^2(1) = 1.86$, $p = .173$, while significantly more participants made a commission error on the first target in the control zero-target modified

¹⁷ Across control conditions, 96.36% of commission errors occurred on the presentation of the no-longer-relevant target word, 2.73% occurred on the subsequent trial, and .91% commission errors occurred on the trial after the subsequent trial. Across lure conditions, 80% of commission errors occurred on the presentation of the first “target” word in the finished PM phase, while 20% occurred on the subsequent trial.

¹⁸ Comparing Experiment 1a (zero-target condition) with Experiment 2 (zero-target modified condition) revealed that there was not a significant difference in the number of participants that made a commission error in the finished PM phase, $\chi^2(1) = 2.28$, $p = .131$.

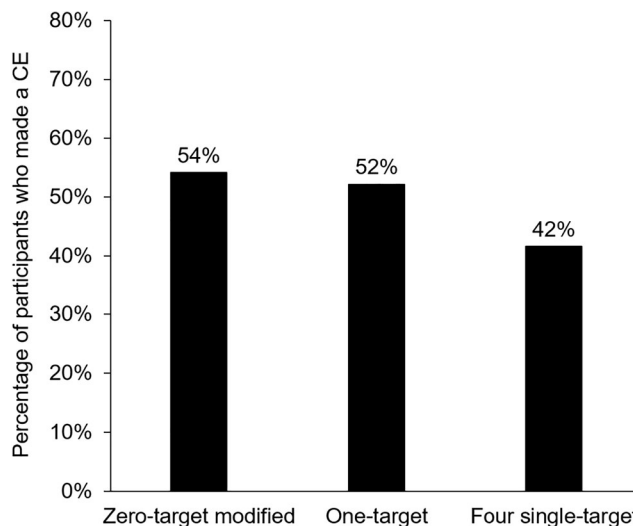


Fig. 5 Percentage of participants who made at least one commission error by condition in Experiment 2

condition (46%) compared with its lure counterpart (0%), $\chi^2(1) = 14.27$, $p < .001$. Additionally, significantly more participants made a commission error in the control one-target condition (48%) compared with its lure counterpart (4%), $\chi^2(1) = 11.78$, $p = .001$ (this difference was not significant, $p = .066$, in Experiment 1b; see Fig. 6).

Lexical decision task performance We examined RTs during the active PM phase and the finished PM phase by performing a 2 (phase: active PM phase, finished PM phase) \times 3 (number of targets: zero, one, four) \times 2 (first target type condition: lure,

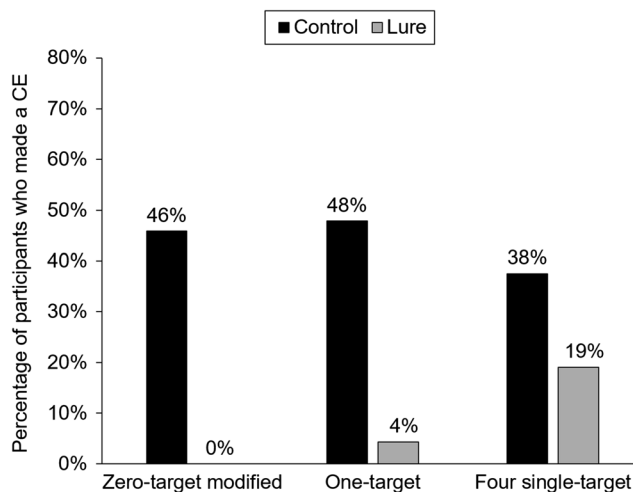


Fig. 6 Percentage of participants who made a commission error (CE) on the first “target” presented in the finished PM phase by condition in Experiment 2. The first “target” in the control conditions was in fact a target on a salient background; first “target” in the lure conditions was a lure trial (nontarget word) on a salient background

control) mixed-model ANOVA (see Table 1). There was a significant effect of phase, $F(1, 134) = 350.72, p < .001$, indicating speeding from the active PM phase ($M = 698, SD = 114$) to the finished PM phase ($M = 549, SD = 79$). No other main effects or interactions were significant. These patterns suggest participants sped up from the active PM phase to the finished PM phase comparably across all conditions.

As for average accuracy (see Table 2), a 2 (phase: active PM phase, finished PM phase) \times 3 (number of targets: zero, one, four) \times 2 (first target type condition: lure, control) mixed-model ANOVA showed a significant effect of phase, $F(1, 134) = 173.70, p < .001$, but no significant effect of number of targets, $F(2, 134) = .96, p = .386$, or first target type condition, $F(1, 134) = 1.69, p = .195$. None of the interactions were significant. Participants performed worse on the ongoing lexical decision task in the active PM phase ($M = 84.17\%, SD = 7.34\%$) than in the finished PM phase ($M = 92.54\%, SD = 5.65\%$).

Combined analysis of Experiments 1a, 1b, and 2

As a reviewer pointed out, a sample size of 24 per cell may yield an underpowered test of some of the comparisons of theoretical interest, and a few of our cells were actually smaller in Experiment 2 due to the interruption of data collection. Some of the comparisons of theoretical interest were significant in one experiment, but not the other (though consistently in the same direction). For example, the contrast in commission error rates between the one-target condition and four single-target condition in Experiment 1a revealed a significantly lower rate for the four single-target condition, whereas this difference was not significant in Experiment 2. Similarly, the contrast in commission error rates between the one-target condition and lure one-target condition in the cross-experimental contrast between Experiments 1a and 1b was not significant, whereas this difference was significant in Experiment 2. To provide a higher-powered test of these and the remaining contrasts (which are all between subjects), we combined the data from all experiments and conducted the commission error analyses reported in the individual experiments. Note, however, that the zero-target conditions were not included in this combined analysis, since we modified the zero-target condition in Experiment 2.

Finished PM phase commission errors

Control conditions (effects of response frequency manipulation) The combined percentage of participants who made a commission error in the one target conditions from Experiments 1a and 2 was compared with the combined percentage of participants who made a commission error in the four

single-target conditions from Experiments 1a and 2. In this combined analysis, significantly fewer participants made a commission error in the four single-target condition (31%) compared with the one-target condition (51%), $\chi^2(1) = 3.85, p = .050$.

For the first no-longer relevant target, fewer participants made a commission error in the four single-target condition (29%) compared with the one-target condition (47%), but this difference was not significant, $\chi^2(1) = 3.14, p = .076$.

Lure conditions (effects of lure manipulation) As a reminder, the primary variable of interest for the lure conditions comparisons was the number of participants who made a commission error on the first “target” that was either a lure (lure condition) or the no-longer-relevant target word (control conditions; see Fig. 7). The combined percentage of participants who made a commission error on the first “target” in the lure one-target conditions from Experiments 1b and 2 was compared with the combined percentage of participants who made a commission error on the first “target” in the control one-target conditions from Experiments 1a and 2. The same comparison was performed for the four single-target conditions. For the one-target conditions, significantly more participants made a commission error on the first target in the control condition (47%) compared with its lure counterpart (13%), $\chi^2(1) = 13.45, p < .001$. In contrast, for the four single-target conditions, the number of participants who made a commission error did not differ between the control (29%) and lure condition (22%), $\chi^2(1) = .59, p = .444$.

Discussion

The primary purpose of Experiment 2 was to attempt to replicate or reproduce (zero-target modified) the patterns

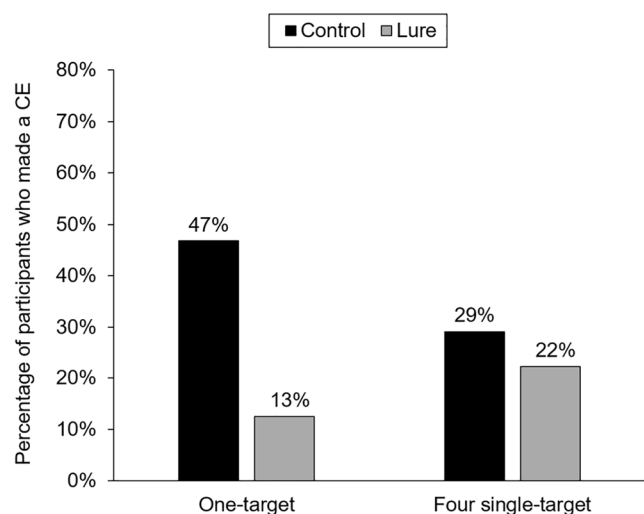


Fig. 7 Percentage of participants who made a commission error (CE) on the first “target” presented in the finished PM phase by condition in the combined analysis

previously observed in the cross-experimental contrast between Experiments 1a and 1b. Those patterns were closely replicated—there was a significantly higher commission error rate for the control zero-target modified and one-target conditions compared with the lure zero-target modified and one-target conditions, respectively, but there was not a difference between the control and lure four single-target conditions. Experiment 2 also enabled us to attempt to replicate the patterns observed within Experiment 1a—namely, the higher commission error rate for the one-target condition compared with the four single-target condition, which had supported the response frequency account. While the rate was again higher for the one-target condition, it was not statistically higher in Experiment 2.

To further test the key theoretical patterns of interest, we combined the data from the one-target and four-single target conditions (control and lure) across all experiments, which enabled higher-powered tests. Not surprisingly, given the results of the individual experiments, the contrast between the control and lure conditions revealed a significant difference for the one-target conditions, with more commission errors in the control condition compared with the lure condition, while the contrast between the control and lure conditions did not differ for the four single-target conditions. Regarding the higher rate of commission errors for the one-target compared with the four single-target condition in Experiment 1a, a difference that was not significant in Experiment 2, the combined analysis revealed a significant difference between these two conditions. We will discuss these findings in more depth in the General Discussion.

A final purpose of Experiment 2 was to investigate whether the rate of commission errors in the zero-target condition might be lower if the finished instructions were revised to eliminate the word *again*. In Experiment 1a, the rate in the zero-target condition was 75%, whereas in the present experiment it was 54%. Although the difference was not statistically significant (see Footnote 18), the reduction in Experiment 2 supports the possibility that some participants who made commission errors in that condition in Experiment 1a may have been inclined to do so because they thought the experimenter made an error (see General Discussion for further discussion). Interestingly, the low(er) rate in the zero-target modified condition was still significantly higher than that of the corresponding lure zero-target modified condition in Experiment 2, further reinforcing the stability of that contrast (lure vs. nonlure).

General discussion

The overarching aim of the current research was to better understand the role that the stimulus–response link plays in intention deactivation and commission errors. Prior work has indicated that intention deactivation plays an important role in

commission error risk but the process by which PM intentions become deactivated has been less clear. Different theoretical accounts of intention deactivation have been proposed, but prior work has not directly examined which account better explains how PM intentions are deactivated. One aim was to address this question by comparing two previously proposed accounts of intention deactivation, the Zeigarnik account and the episodic retrieval account, by testing two hypotheses that fell out of these accounts termed the degree of fulfillment hypothesis and the response frequency hypothesis, respectively. The degree of fulfillment hypothesis suggests commission errors occur due to perseveration of PM intentions that have not been fulfilled. In contrast, the response frequency account posits that performing the PM intention more frequently in the active PM phase creates a stronger stimulus–response link that allows a stop tag to be connected to the link, making it easier to deactivate the PM intention.

The evidence was somewhat mixed with respect to these hypotheses. Experiment 1a supported the response frequency account in demonstrating that rates of commission errors differed between two conditions that were matched on the degree of intention fulfillment (i.e., the one-target and four single-target conditions), but varied with respect to the number of responses that were made to the presented target. Responding to the target multiple times (four single-target condition), which presumably strengthened the stimulus–response link for that target, led to fewer commission errors than responding just once (one-target condition). While Experiment 2 again found lower rates of commission errors for the four single-target condition compared with the one-target condition, the difference was not significant. Finally, combining the data from Experiments 1 and 2 to produce a higher-powered test, the commission error rate was significantly lower for the four single-target condition than the one-target condition. Collectively, the findings indicate an effect of the response frequency manipulation on commission error risk that favors the response frequency account; however, the effect was not stable across the two subsamples (Experiments 1a and 2), and therefore we cannot fully rule out the degree of fulfillment account.

While Experiment 1a examined the role that the stimulus–response link plays in intention deactivation, Experiment 1b focused on what exactly the “stimulus” in the stimulus–response link is. Although one might assume that participants link their PM response to the target words, it is possible that their responses are being linked to the context in which the words appear (e.g., the salient background). To examine this possibility, we embedded lure trials in the finished PM phase in Experiment 1b. Lure trials matched the salient background associated with target trials, but contained a different word (e.g., if targets were *corn* and *dancer*, the lure was *fish*). We then compared commission error rates on the lure trials to the control (target) trials from Experiment 1a. Furthermore, we directly contrasted these conditions head-to-head within a

single experiment in Experiment 2. Providing evidence that the stimulus itself (the target word) and not solely the context (salient background) was linked to the PM response, participants in the zero-target and one-target conditions were less likely to make a commission error on lure trials compared with target trials. This is an important finding because it suggests that, under these conditions, context alone may not be enough to cause commission errors. However, in contrast, in the four single-target condition, participants were just as likely to make a commission error to a lure trial as a target trial, suggesting reliance on context in this case. Notably, these patterns were consistent across the various experiments (Experiments 1a vs. 1b, Experiment 2) and in the combined analysis, suggesting a stable and robust effect of the lure manipulation on commission error risk.

The findings in Experiment 1b stimulate an interesting question: Why do participants appear to link the target word to the response in some conditions, but not in other conditions? One possible explanation is due to the predictive nature of the salient context during the active PM phase. Let us first consider the four single-target condition where participants encountered four target word trials and each one appeared in the salient context. It is possible that participants' representation of the "stimulus" in the four single-target condition *initially* may have consisted of a combination of the target word and the salient background (similar to the other conditions) given the initial instructions during encoding; however, as the predictive value of the salient context increased (i.e., the salient context always correctly predicted the presence of a target word, and the salient context and target word repeatedly appeared together in the active PM phase), participants' representation of the "stimulus" may have shifted to the salient context alone. In most circumstances, this shift would be both logical and efficient, but in this experiment, it may have made participants in the four single-target conditions susceptible to commission errors on the first target in the finished PM phase, including when that target was a lure. By contrast, participants in the one-target condition encountered only one target word in the active PM phase, and although that word appeared in the salient context, a single experience with this pairing may not have been sufficient for participants to associate the context alone with the intention (i.e., to rely on the context to trigger intention retrieval). In the case of the zero-target condition, participants did not encounter any targets and thus never encountered the salient context during the active PM phase. On our view that the context is relied upon to the extent that it is predictive of target occurrence, it is unsurprising that the zero-target condition maintained a representation of the target word (possibly in conjunction with the context) as opposed to relying solely on context.

The effects of the lure manipulation are also interesting when considered from the perspective of the dual-mechanism account of commission errors. This account has

posited a role for the salient context in stimulating spontaneous retrieval of PM intentions in the context of the traditional four-target condition (Scullin & Bugg, 2013; Scullin et al., 2012; Scullin et al., 2011). However, these prior studies did not include lure trials and could not determine if the salient context alone could lead to commission errors.¹⁹ At least in the current four single-target condition, it appears this is possible. However, in the one-target and zero-target conditions, this is not the case. The implication is that spontaneous retrieval of the intention in these latter two conditions may stem from processing either the target word alone or the conjunction of the target word and the salient context.

Limitations and future directions

Although the findings from Experiment 1a and the combined analysis favored the episodic trace account over the Zeigarnik account, given that the contrast between the four single-target and one target conditions in Experiment 2 supported the degree of fulfillment hypothesis, it is important for future research to continue to examine both accounts. It is likely that the response frequency hypothesis has boundaries. For example, Pink and Dodson (2013) showed that a PM intention that was performed 10 times for each of eight targets became habitual and was difficult to deactivate. In the current study, the maximum number of times that a PM target was responded to was four. Examining a fuller range of possible responses may elucidate the function relating response frequency to intention deactivation. It is also plausible that there may be some contexts in which perseveration of an unfulfilled or partially fulfilled PM intention may lead to a commission error, providing further support for the degree of fulfillment hypothesis. For example, future research might examine whether participants are more likely to make a commission error in the current four single-target condition compared with the traditional four-target condition. If so, this would support the degree of fulfillment hypothesis because the current condition is a partial intention fulfillment condition whereas the traditional condition is a complete fulfillment condition. Furthermore, although the total number of responses (four) is equated across these two conditions, in the traditional condition participants respond twice to a given target whereas in the current condition, they

¹⁹ These prior studies included control trials. For example, in the traditional four-target condition, if the target words appeared on a red background on four trials, then control words appeared on a blue background on four trials. Participants rarely false alarmed to the control trials (active PM phase) and rarely made commission errors in response to control trials (finished PM phase), suggesting that the triggering feature was potentially a specific salient context (e.g., red screen), but was not merely the context change (change from a black screen to a different colored screen). Similarly, in the repeated cycles paradigm where slowing is observed on previously irrelevant targets compared with oddballs (Walser et al., 2012), which also differ from standard ongoing trials on one visual feature, it can be assumed that not merely the deviation itself is the triggering feature as otherwise there should not be slowing.

respond four times to a given target. Thus, a response frequency account may, if anything, predict the opposite pattern—higher commission error risk in the traditional four-target condition.

As for our findings regarding the effects of the lure manipulation, while we can conclude that participants in the zero-target and one-target conditions were not linking their PM intention to the salient context alone, a limitation of our design is that we are not able to differentiate between the other two possible “stimulus” representations: the target word alone or a combination of the target word and the salient context. One clear prediction that falls out of the explanation we forwarded above based on the predictive nature of the salient context is that reliance exclusively on the target word should increase to the extent that one reduces the predictability of the target based on the context (e.g., if the target appeared in a unique context each time). Testing this prediction may prove challenging, however, as prior research has shown that participants are unlikely to commit commission errors unless they appear in a salient (and predictive, to date) context (Scullin et al., 2012; Scullin et al., 2011).

Finally, the findings of Experiment 2 should motivate additional research to further explore the sources of commission errors in the zero-target condition. At the start of this research, we considered that the high rates of commission errors found in this condition may reflect either the absence of intention fulfillment (creating intention perseveration) or the absence of prior responses (and therefore, representations to bind a stop tag to). The finding that commission error rates were nominally lower in the modified zero-target condition in Experiment 2 raises the possibility that another source is possible—participants may have pressed the *Q* key in the standard variant of this condition (e.g., Experiment 1a) because they thought the experimenter made an error upon being instructed not to perform the task “again.” Considering past research that has demonstrated higher rates of commission errors in zero-target conditions that did not use phrasing such as the word *again* (see, e.g., Schaper & Grundgeiger, 2017, who told participants they could ignore the red screen in the finished phase because they were chosen for a condition that did not have to react to it) we also cannot rule out the possibility that the Experiment 2 commission error rate reflects sampling error. Thus, future studies should contrast the standard and modified versions head-to-head to determine the stability of this pattern.

Conclusion

Collectively, the findings provided initial evidence demonstrating the importance of different aspects of the stimulus–response link to intention deactivation and accordingly, commission error risk. Consistent with the episodic trace account, one take home message is that prior responding appears to

facilitate later deactivation of an intention and the benefits of prior responding can be distinguished from those of intention fulfillment. However, while we provided evidence in Experiment 1a and the combined analysis for the response frequency hypothesis, Experiment 2 did not replicate that pattern, and therefore we cannot yet conclude that it is robust. The second take-home message is that what precisely is stored and associated in the episodic traces of prior responding is not always a target–response link; rather, in some cases, responses are bound merely to contextual cues that are predictive of target occurrence as revealed by the novel lure manipulation. The effect of the lure manipulation appears to be robust as the evidence supporting this conclusion was strong and consistent across experiments and the combined analysis. These novel findings bring us another step closer to understanding the processes underlying intention deactivation. Future research should aim to further test theoretical accounts of intention deactivation and evaluate applications of this knowledge to prevent commission errors from occurring in real-life settings.

Acknowledgements Both authors were supported by National Institute on Aging, Grant R21AG052060. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging. The authors are grateful to Casey Bowen, Megan Maxwell, Nancy Yang, and Hannah Yoken for assistance with data collection.

Open practices statement None of the data or materials for the experiments reported here is publicly available, and none of the experiments was preregistered.

References

- Anderson, F. T., & Einstein, G. O. (2017). The fate of completed intentions. *Memory*, 25(4), 467–480. doi:<https://doi.org/10.1080/09658211.2016.1187756>
- Bugg, J. M., & Ball, B. H. (2017). The strategic control of prospective memory monitoring in response to complex and probabilistic contextual cues. *Memory & Cognition*, 45(5), 755–775. doi:<https://doi.org/10.3758/s13421-017-0696-1>
- Bugg, J. M., & Scullin, M. K. (2013). Controlling intentions: The surprising ease of stopping after going relative to stopping after never having gone. *Psychological Science*, 24(12), 2463–2471. doi:<https://doi.org/10.1177/0956797613494850>
- Bugg, J. M., Scullin, M. K., & Rauvola, R. S. (2016). Forgetting no-longer-relevant prospective memory intentions is (sometimes) harder with age but easier with forgetting practice. *Psychology and Aging*, 31(4), 358–369. doi:<https://doi.org/10.1037/pag0000087>
- Bugg, J. M., & Streeper, E. (2019). Fate of suspended and completed prospective memory intentions. In M. A. McDaniel & J. Rummel (Eds.), *Current issues in memory: Prospective memory* (pp. 43–58). New York, NY: Taylor & Francis.
- Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 717–726. doi:<https://doi.org/10.1037/0278-7393.16.4.717>
- Einstein, G. O., McDaniel, M. A., Smith, R. E., & Shaw, P. (1998). Habitual prospective memory and aging: Remembering intentions

- and forgetting actions. *Psychological Science*, 9(4), 284–288. doi:<https://doi.org/10.1111/1467-9280.00056>
- Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus–response episodes. *Visual Cognition*, 5(1/2), 183–216. doi:<https://doi.org/10.1080/713756773>
- Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral & Brain Sciences*, 24(5), 849–878. doi:<https://doi.org/10.1017/S0140525X01000103>
- Lourenço, J. S., White, K., & Maylor, E. A. (2013). Target context specification can reduce costs in nonfocal prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(6), 1757–1764. doi:<https://doi.org/10.1037/a0033702>
- Marsh, R. L., Hicks, J. L., & Bink, M. L. (1998). Activation of completed, uncompleted, and partially completed intentions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(2), 350–361. doi:<https://doi.org/10.1037/0278-7393.24.2.350>
- McDaniel, M. A., Bugg, J. M., Ramuschkat, G. M., Kliegel, M., & Einstein, G. O. (2009). Repetition errors in habitual prospective memory: Elimination of age differences via complex actions or appropriate resource allocation. *Aging, Neuropsychology, and Cognition*, 16(5), 563–588. doi:<https://doi.org/10.1080/13825580902866646>
- Möschl, M., Fischer, R., Bugg, J. M., Scullin, M. K., Goschke, T., & Walser, M. (2020). Aftereffects and deactivation of completed prospective memory intentions: A review. *Psychological Bulletin*, 146, 245–278.
- Pink, J. E., & Dodson, C. S. (2013). Negative prospective memory: Remembering not to perform an action. *Psychonomic Bulletin & Review*, 20(1), 184–190. doi:<https://doi.org/10.3758/s13423-012-0337-4>
- Schaper, P., & Grundgeiger, T. (2017). Commission errors in delay-execute prospective memory tasks. *The Quarterly Journal of Experimental Psychology*, 70(8), 1423–1438. doi:<https://doi.org/10.1080/17470218.2016.1187182>
- Scullin, M. K., & Bugg, J. M. (2013). Failing to forget: Prospective memory commission errors can result from spontaneous retrieval and impaired executive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 965–971. doi:<https://doi.org/10.1037/a0029198>
- Scullin, M. K., Bugg, J. M., & McDaniel, M. A. (2012). Whoops, I did it again: Commission errors in prospective memory. *Psychology and Aging*, 27(1), 46–53. doi:<https://doi.org/10.1037/a0026112>
- Scullin, M. K., Bugg, J. M., McDaniel, M. A., & Einstein, G. O. (2011). Prospective memory and aging: Preserved spontaneous retrieval, but impaired deactivation, in older adults. *Memory & Cognition*, 39(7), 1232–1240. doi:<https://doi.org/10.3758/s13421-011-0106-z>
- Walser, M., Fischer, R., & Goschke, T. (2012). The failure of deactivating intentions: Aftereffects of completed intentions in the repeated prospective memory cue paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(4), 1030–1044. doi:<https://doi.org/10.1037/a0027000>
- Zeigarnik, B. (1938). On finished and unfinished tasks. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 300–314). London, England: Kegan Paul, Trench, Trubner & Company. doi:<https://doi.org/10.1037/11496-025>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.