TRANSFERABLE INHIBITION OF DIRECT SUPPRESSION: EVIDENCE FROM A DOT-PROBE TASK

Satoru NISHIYAMA[1), 2)] and Satoru SAITO[1)]

[1)]*Kyoto University, Japan*
[2)]*Japan Society for the Promotion of Science, Japan*

Previous studies on the Think/No-Think (TNT) paradigm have demonstrated that retrieval stopping causes later forgetting. Although precise mechanisms of this retrieval stopping effect have come under scrutiny, a recent study (Hertel & Hayes, 2015) has provided a signpost finding; that is, in a flanker task subsequent to a TNT task, ratings of words flanked by cues for retrieval stopping were slower among those who stopped retrieval by pushing memories out of awareness (i.e., direct suppression), but not among those who did so by thinking of another memory (i.e., thought substitution). This result is attributable to two possible mechanisms: cues for direct suppression triggered some inhibitory mechanisms (inhibition transfer) or they drew more attention than other cues (attentional capture). The present study clarifies the aftereffect of direct suppression by conducting a TNT experiment with a dot-probe task, wherein probes appear following TNT cues, including suppression ones. Participants who engaged in thought substitution showed attentional capture by TNT cues; faster responses to probes presented in the same position of TNT cues, and slower responses to probes presented in the opposite position. However, those who engaged in direct suppression did not show attentional capture. These results are clearly inconsistent with the attentional capture account for the aftereffect of direct suppression, which, in turn, favours inhibition transfer account. Correlation analyses also supported inhibition transfer. Repeated direct suppression might associate suppression cues with inhibitory mechanisms, resulting in inhibition exerted automatically by the presentation of suppression cues.

Key words: Think/No-Think paradigm, forgetting, attention, inhibition, dot-probe task

Various memories come to mind in our daily lives. We can usually think about some memories while avoiding others, depending on the situation. This mental activity is referred to as memory control and is important not only for task performance but also for mental health (Nørby, 2015, 2018). Maintaining memories relevant to the task at hand contributes to better performance. In addition, stopping the retrieval of distressing memories can be beneficial for one's psychological well-being.

Memory control, especially retrieval stopping, has been studied using the Think/No-Think (TNT) paradigm (Anderson & Green, 2001). In the TNT paradigm, participants first learn item pairs. Then, participants perform the critical TNT phase, in which one item in the pair is presented as a cue, and participants are required to retrieve (Think) or stop retrieving (No-Think) the corresponding item (target). Subsequently, participants' memories are tested by asking them to recall the targets from the cues. TNT studies have shown enhanced recall of Think targets and impaired recall of No-Think targets compared to control counterparts (i.e., Baseline targets), which are studied before the TNT phase, but are neither retrieved nor avoided during the TNT phase.

Recently, Hertel and Hayes (2015) found that the effect of retrieval stopping carries over to a task with little relation to memory. In their two experiments, the TNT phase was followed by a flanker task, which is normally thought to require response inhibition (Friedman & Miyake, 2004). In that flanker task, participants rated the valence of a word that was presented in the centre of the screen. At the same time, another word was presented as a flanker below and above the word to be rated. For some trials, cues for TNT pairs were presented as flankers, and the mean reaction times (RTs) for rating words flanked by different types of TNT cues (i.e., Think, No-Think, and Baseline) were contrasted. The results showed that participants took longer to rate words when flankers were No-Think cues than when flankers were Think or Baseline cues. Moreover, the rating delay was positively correlated with the degree of memory impairment. Importantly, these results were observed among individuals who were engaged in direct suppression when stopping retrieval but not among participants who were engaged in thought substitution. Direct suppression and thought substitution are two distinct strategies for retrieval stopping. Direct suppression is a strategy to keep some memories away from one's awareness without engaging in any distracting thoughts. Conversely, thought substitution encourages one to use irrelevant thoughts to avoid the retrieval of unwanted memories. These strategies utilize different mechanisms. Specifically, direct suppression involves inhibitory processes more strongly than thought substitution (Benoit & Anderson, 2012; Hulbert et al., 2016; Wang et al., 2015). Based on performance differences between direct suppression and thought substitution groups, Hertel and Hayes (2015) concluded that the inhibitory reactions acquired through repeated suppression activities transferred to the unrelated task via No-Think cues. That is, cues for retrieval stopping based mainly on inhibitory mechanisms unintentionally trigger inhibitory reactions in processes for other unrelated tasks (in this case, a valence rating).

This inhibition transfer account is plausible because recent evidence indicates that a key component of neural networks is shared across retrieval stopping, emotion regulation, and motor inhibition (Depue et al., 2016). However, rating delays can also be interpreted in terms of spatial attention, namely, the attentional capture account. During the TNT phase, participants are required to focus entirely on the presented cues when stopping retrieval (i.e., No-Think trials) in the direct suppression strategy. In contrast, participants do not need to focus on cues when retrieving (i.e., Think trials) or when stopping retrieval using a thought substitution strategy; instead, participants have to be engaged in the paired targets or other irrelevant thoughts, respectively. This would result in more

attention allocated to No-Think cues used during direct suppression than to Think and Baseline cues, as well as to No-Think cues used during thought substitution. This might be the reason why participants engaged in direct suppression took longer to rate the valence of words flanked by No-Think cues. No-Think cues distracted participants' attention away from the words to be rated. This account is consistent with the positive correlation between memory impairment and rating delays observed in the direct suppression group. Here, individuals who adequately inhibited their memories were more engaged with the No-Think cues to stop retrieval, resulting in greater attention captured by those No-Think cues.

These two accounts predict distinct consequences of using a direct suppression strategy for forgetting unwanted memories in daily life. The inhibition transfer account predicts beneficial outcomes of direct suppression, but the attentional capture account predicts ineffective, even harmful, outcomes. In terms of the inhibition transfer account, after direct suppression, cues for suppressed memories work as cues for inhibitory mechanisms; thus, suppressed memories are less likely to be retrieved when one encounters these cues again. In contrast, in terms of the attentional capture account, suppressed memories are less likely to be retrieved at first, but the effect declines because cues draw more attention and increase the probability of the retrieval of unwanted memories. Thus, it is theoretically and practically important to evaluate which account is a more plausible explanation for the aftereffect of direct suppression.

For this purpose, the current study employed a dot-probe task (MacLeod et al., 1986) in which participants were required to respond to a probe presented subsequent to a pair of prime stimuli. Primes comprised a TNT cue and a novel word during critical trials. Importantly, positions of the prime TNT cue and probe determined whether a trial was congruent or incongruent. For congruent trials, probes were presented in the same position as the TNT cue, while the trials in which probes were presented to the opposite side of a TNT cue (i.e., the position of the novel word) were denoted as incongruent trials. Differences in RTs across the two trial types (congruency effect) are considered a measure of attentional preference. Importantly, the inhibition transfer and attentional capture accounts provide different predictions regarding RTs for trials where primes contain No-Think cues. The inhibition transfer account predicts that responses to probes would be slower during both types of trials in which primes contain No-Think cues than in the other trials; this is because responses are inhibited by certain mechanisms triggered by the No-Think cues. In contrast, the attentional capture account predicts that the congruency effect would be greater when one of the primes is a No-Think cue compared to when a prime is a Think or Baseline cue; this is because the attentional capture account assumes that No-Think cues would draw more attention than Think or Baseline cues. No-Think cues guide participants' attention to their position, resulting in a faster response to the probe at the same position (congruent trials) but a slower response to the probe at the opposite position. The two accounts would also predict different correlations between memory impairment and performance on the dot-probe task. The inhibition transfer account predicts a positive correlation between memory impairment and RTs during the dot-probe task; more specifically, individuals who are better able to suppress

memories should be slower to respond to the probes. Conversely, the attentional capture account predicts a negative correlation between memory impairment and RTs during congruent trials but a positive correlation between memory impairment and RTs during incongruent trials. This is because individuals who are better able to suppress memories should pay more attention to No-Think cues.

Although the aforementioned predictions are concerned with direct suppression, a thought substitution group was also included in the present study to confirm that thought substitution induces neither inhibition transfer nor attentional capture, as demonstrated in Hertel and Hayes (2015). Specifically, it was predicted that participants in the thought substitution group would show no differences in overall RT and the congruency effect between any prime trial conditions. Thus, RTs to a probe following a prime containing No-Think cues would be slower in the direct suppression group than in the thought substitution group. In line with the prediction that thought substitution causes neither inhibition transfer nor attentional capture, there should be no relationship between memory impairment and RTs during the dot-probe task for the thought substitution group. In Experiment 1, a basic TNT finding—memory impairment via retrieval stopping—was assessed to confirm that our design was appropriate for inducing forgetting. In Experiment 2, the dot-probe task was conducted after the TNT phase to examine the main hypotheses. The Research Ethics Review Board for Experimental Psychology at the Graduate School of Education, Kyoto University, approved this study.

EXPERIMENT 1

METHOD

Participants

Thirty-six undergraduate and graduate students (19 men, mean age = 21 ± 1.3 years) participated in this experiment. This sample size was the same as the sample size used in Benoit and Anderson (2012). Participants were native Japanese speakers with no history of mental disorders. Participants were recruited through an online website housed by the Kyoto University Co-op and received a book coupon for their participation. Written informed consent was obtained at the beginning of the experiment.

Design

Experiment 1 was conducted with a two-factor mixed design. The first factor, No-Think strategy, had 2 levels (direct suppression and thought substitution), which were manipulated between participants. The second factor, TNT status, had 3 levels (Think, No-Think, and Baseline), which were manipulated within participants.

Materials

Thirty-six word triplets, each comprising a cue, target, and substitute, were used. All words were selected from Japanese nouns with two to five morae and high familiarity (more than 5.5 of 7 points) in a Japanese corpus (Amano & Kondo, 1999). During selection, written forms were taken into consideration (*Kanji*, *Hiragana*, and *Katakana*). These forms could affect performance (especially depending on one's visual acuity) in the TNT task as well as the dot-probe task used in Experiment 2. Cues were written with two *Kanji* (i.e., Chinese) characters, while targets and substitutes were written in *Hiragana* or *Katakana* (i.e., Japanese phonograms). Within a triplet, a cue and target had a weak association, while a cue and

substitute had a strong association. There was, however, no association between the triplets. Ten filler triplets were created in the same way.

Procedure

Experiment 1 comprised four phases: encoding, TNT task, recall test, and post-experiment questionnaire. During the encoding phase, participants learned all cue-target pairs, including filler pairs, in the following three steps. First, cue-target pairs were presented for 5 s. Then, participants performed memory tests with feedback. Participants were provided with cues and asked to recall the corresponding targets within 5 s. After participants responded, or 5 s passed without a response, the correct target was presented for 2 s as feedback. The experimenter recorded whether the participants' responses were correct by pressing a key. Participants repeated this test-and-feedback cycle until they could correctly recall more than half of the targets. Finally, participants took a memory test (hereinafter referred to as the initial learning test) in the same way as the test-and-feedback cycles; however, this time, participants were allowed 3 s to recall, and no feedback was provided.

During the TNT phase, participants were required to think or not to think about the targets associated with the presented cues. When cues were presented in green, participants were instructed to think of the corresponding targets, whereas when cues were presented in red, participants were instructed to stop retrieval of the corresponding targets. To stop retrieval, participants were provided two instructions: direct suppression and thought substitution. Direct suppression involved focusing on the cues and pushing away any other thoughts, including the targets, from awareness. Thought substitution involved preventing targets from coming to mind by thinking about substitutes, which were studied after the task instructions were given. To ensure a distinction between the two instructions, participants practiced the TNT task three times. During the first practice, participants performed No-Think trials following the direct suppression instruction. Next, after studying substitutes for the No-Think targets, which were presented for 2 s, participants performed the second TNT practice in which they engaged in thought substitution. Finally, participants were assigned to either strategy condition and performed the third TNT practice with the assigned strategy. Across all practice conditions, cues for the filler sets were presented for 3 s.

After participants studied the substitutes again, they carried out the main TNT task in the same way that they did during the final practice task. The main TNT task comprised 288 trials, which were divided into 12 blocks. Each block contained 12 Think and 12 No-Think trials. Twenty-four cues assigned to either the Think or No-Think condition were presented once in each block. Participants took a break for 60 s every three blocks. After the sixth block was completed, participants studied the substitutes again, regardless of the No-Think strategy used.

After the TNT task, two cued recall tests were presented to assess participants' memories of the targets and substitutes. During the first cued recall test, participants were given all cues, except for filler cues, and were required to recall the corresponding targets. During the second cued recall test, participants were given only 12 No-Think cues and were required to recall the corresponding substitutes. Cues were presented for 3 s during both tests.

At the end of the experiment, participants completed a thought evaluation questionnaire (hereinafter referred to as the "questionnaire") that listed Think and No-Think cues and the corresponding target and substitutes with five-point scales. Participants were instructed to rate how much they thought about a cue, target, and substitute when each cue was presented during the TNT phase (1 = *not at all* to 5 = *very much*). Participants also rated the substitutes for the Think targets that were prepared for counterbalancing but not presented during the experiment. This questionnaire was intended to assess whether participants performed the TNT task as instructed. Though the questionnaire did not directly ask participants which strategy they used during the No-Think trials, we considered responses to be a measure of strategy use, as completely different ratings would likely be obtained depending on the strategy. Participants in the direct suppression group would likely indicate high scores for No-Think cues and low scores for the corresponding targets and substitutes. In contrast, participants in the thought substitution group would likely rate No-Think cues and substitutes with high scores, but this would not be the case for targets. It was also predicted that participants in both No-Think strategy groups would rate Think item sets in the same way. More specifically, participants likely would rate Think cues and corresponding targets with high scores but not the substitutes for Think items. All procedures, except the questionnaire, were implemented using PsychoPy 1.82.01 (Peirce, 2007, 2009).

RESULTS

We conducted all statistical analyses in R 4.0.2 (R Core Team, 2020) with the package *rstatix* 0.6.0 (Kassambara, 2020). All data and R scripts are available at GitHub (https://github.com/snishiyama/TNT-DP).

Thought Evaluation Questionnaire Ratings

Table 1 presents the mean scores for all ratings. All scores were consistent with predictions, which confirmed that participants performed the TNT task as instructed. A series of 2 (No-Think strategy) × 3 (word type: cue, target, and substitute) analyses of variance (ANOVAs) were conducted separately for scores in the Think and No-Think conditions. For the Think condition, only a main effect of item type was significant (F $(1.44, 47.97) = 176.39$, $p < .001$, $\eta_G^2 = 0.77$).[1] Post hoc t-tests[2] showed that both cues and targets were scored higher compared with substitutes (cues vs. substitutes, $t(34) = 13.14$, $p < .001$, Hedges' $g = 2.96$; targets vs. substitutes, $t(34) = 28.85$, $p < .001$, Hedges' $g = 7.20$). For the No-Think condition, both main effects and the interaction were significant (No-Think strategy, $F(1, 34) = 7.55$, $p = .010$, $\eta_G^2 = 0.10$; word type, F $(2, 68) = 19.11$, $p < .001$, $\eta_G^2 = 0.21$; interaction, $F(2, 68) = 29.12$, $p < .001$, $\eta_G^2 = 0.29$).

Table 1.   Mean and 95% CI for the Thought Evaluation Questionnaire Ratings in Experiments 1 and 2

		Direct suppression			Thought substitution		
		Cue	Target	Substitute	Cue	Target	Substitute
		Experiment 1					
Think	Mean	3.62	4.24	1.11	3.69	4.40	1.16
	95% CI	[3.08, 4.15]	[3.93, 4.55]	[1.02, 1.20]	[3.05, 4.33]	[4.13, 4.66]	[1.03, 1.29]
No-Think	Mean	4.42	2.36	2.44	3.50	3.34	4.19
	95% CI	[4.15, 4.69]	[1.96, 2.76]	[2.08, 2.80]	[2.84, 4.15]	[2.79, 3.88]	[3.83, 4.54]
		Experiment 2					
Think	Mean	3.84	4.00	1.19	3.70	4.45	1.20
	95% CI	[3.38, 4.30]	[3.61, 4.38]	[1.08, 1.29]	[3.15, 4.24]	[4.22, 4.67]	[1.13, 1.28]
No-Think	Mean	4.46	2.43	2.37	3.74	2.62	4.23
	95% CI	[4.20, 4.72]	[2.11, 2.74]	[2.05, 2.70]	[3.20, 4.29]	[2.31, 2.93]	[3.90, 4.56]

Note. CI = Confidence interval.

---

[1] The degrees of freedom for the F-statistic were adjusted by the Greenhouse-Geisser correction as Mauchly's Test of Sphericity indicated that the sphericity assumption had been violated ($p < .05$). This correction was applied to all the repeated measure (or mixed) ANOVA throughout this article.
[2] p-values were adjusted by the Holm method. This adjustment was applied to all the post hoc t-tests throughout this article.

Simple main effect analyses showed significant main effects of item type for the two No-Think strategies (direct suppression, $F(1.25, 21.31) = 59.42$, $p < .001$, $\eta_G^2 = 0.66$; thought substitution, $F(2, 34) = 4.83$, $p = .014$, $\eta_G^2 = 0.11$). Post hoc t-tests revealed that in the direct suppression group, cues received higher scores than targets and substitutes (cues vs. targets, $t(17) = 7.95$, $p < .001$, Hedges' $g = 2.95$; cues vs. substitutes, $t(17) = 8.11$, $p < .001$, Hedges' $g = 3.01$). In contrast, in the thought substitution group, substitutes received higher scores than targets ($t(17) = 2.95$, $p = .027$, Hedges' $g = 0.90$).

Recall Rate

The target and substitute recall rates are shown in Table 2. Only learned targets (i.e., targets that were recalled correctly during the initial learning test) were included in the analysis. A 2 (No-Think strategy) × 3 (TNT status) ANOVA revealed a significant main effect of TNT status ($F(1.49, 50.68) = 17.27$, $p < .001$, $\eta_G^2 = 0.17$), but the main effect of No-Think strategy and the interaction were not significant (Fs < 1). As in previous studies, post hoc t-tests showed that the recall rate for No-Think targets was significantly lower than for any other target (No-Think vs. Think, $t(34) = 4.49$, $p < .001$, Hedges' $g = 0.89$; No-Think vs. Baseline, $t(34) = 4.37$, $p < .001$, Hedges' $g = 0.75$).

The pattern of substitute recall was also replicated. Consistent with Benoit and Anderson (2012), Welch's t-test revealed that substitutes were recalled better in the thought substitution group than in the direct suppression group ($t(19.76) = 3.05$, $p = .006$, Hedges' $g = 0.99$).

Table 2.   Mean and 95% CI for Recall Rates in the Final Recall Test in Experiments 1 and 2

		Think	No-Think	Baseline	Substitute
		Experiment 1			
Direct suppression	Mean	0.99	0.89	0.97	0.83
	95% CI	[0.97, 1.01]	[0.82, 0.95]	[0.94, 1.00]	[0.74, 0.92]
Thought substitution	Mean	0.98	0.90	0.97	0.96
	95% CI	[0.95, 1.01]	[0.84, 0.96]	[0.94, 1.01]	[0.94, 0.99]
		Experiment 2			
Direct suppression	Mean	0.99	0.93	0.98	0.87
	95% CI	[0.98, 1.00]	[0.89, 0.97]	[0.97, 1.00]	[0.80, 0.93]
Thought substitution	Mean	1.00	0.91	0.99	0.99
	95% CI	[0.99, 1.00]	[0.85, 0.98]	[0.98, 1.00]	[0.98, 1.00]

Note. CI = Confidence interval.

<div align="center">

DISCUSSION

</div>

Experiment 1 was a direct replication of previous TNT studies (e.g., Benoit & Anderson, 2012). The below-baseline recall of No-Think targets was successfully replicated in both No-Think strategy groups. The substitute recall results and thought evaluation ratings suggest that participants followed the No-Think instructions. Mean ratings were higher for the substitutes than the other cues in the thought substitution group; this difference in score was due to participants not having to focus on cues internally. Hence, the stimuli and design used were appropriate for stimulating retrieval stopping. Therefore, the next experiment was carried out, in which participants performed a dot-probe task after retrieval stopping, to examine whether inhibition transfer or attentional capture is more responsible for the aftereffects of direct suppression.

It should be noted that the experimenter, who sat behind the participant during Experiment 1, observed that participants had difficulty recalling No-Think targets that were finally recalled correctly within the required time; however, no data was collected to fully confirm this observation. Thus, every trial during the recall tests was recorded in Experiment 2 to measure recall latencies and examine whether retrieval stopping would cause not only a reduction in recall but slower recall of No-Think items. According to the experimenter's observation, it was predicted that recall latencies for No-Think items would be longer compared to that of any other target in both strategy groups, even when controlling for recall latencies during the initial recall test.

<div align="center">

EXPERIMENT 2

METHOD

</div>

Participants

Forty-eight undergraduate and graduate students (29 men, mean age = 21 ± 1.6 years) participated. This sample size was determined based on the number of counterbalanced conditions, as well as the sample size used in Benoit and Anderson (2012). Participants were all native Japanese speakers with no history of mental disorders. Participants were recruited through an online website via the Kyoto University Co-op and received a book coupon for their participation. Written informed consent was obtained at the beginning of the experiment. One participant was excluded from analyses because the participant was suspected of not being compliant with the instructions. In the thought evaluation questionnaire, the participant scored higher than two on average for the rating of think cue-associated substitutes although the participant had not seen these substitutes in the task.

Design

Experiment 2 was conducted with a three-factor mixed design. The first and second factors were identical to those in Experiment 1: No-Think strategy and TNT status. The third factor, congruency, had 2 levels (congruent and incongruent), which were manipulated within participants.

Materials

The word triplets in Experiment 1 were used, although a few cues that shared the same *Kanji* with each other were replaced with new ones. Another item set of 160 word pairs was constructed for the dot-probe task. The word pairs comprised 36 words used as TNT cues and 124 new words that were also

selected from Japanese nouns with high familiarity (more than 5.5 of 7 points) in the Japanese corpus (Amano & Kondo, 1999). All words in these sets were written with two *Kanji* characters and did not share the same *Kanji* characters with each other. Each letter of the paired words had nearly the same number of strokes to minimize the influence of word complexity on attention. For counterbalancing, two distinct word pairs were created.

Procedure

Experiment 2 comprised five phases: encoding, TNT task, dot-probe task, recall test, and thought evaluation questionnaire. The TNT procedures (encoding, TNT task, recall test, and questionnaire) were identical to Experiment 1, except for two aspects. First, the test-and-feedback cycle was repeated until participants could correctly recall more than 90% of the targets (or up to three times) to maximize the number of trials available for the dot-probe task and final recall test. Although two participants failed to achieve the 90% recall rate by the third cycle, they were not excluded because these two participants could recall 32 of 36 (approximately 89%) critical targets. Second, verbal responses were recorded during every trial in the initial learning and final recall tests to measure recall latency as an index of retrieval efficiency.

During the dot-probe task, each trial started with a fixation cross presented in the centre of the display for 1 s. Subsequently, a pair of words was presented vertically as prime stimuli for 500 ms, with a distance of 3 cm. Exactly 1/60 s after the primes disappeared, a probe (E or F) appeared in either the upper or lower location where the words had been presented. Participants were asked to identify which probe, E or F, was presented by pressing a given key on a QWERTY keyboard as quickly and accurately as possible. Participants were instructed to press the 'x' key with their left index finger if E was presented or the '.' key with their right index finger if F was presented. As soon as participants responded, the display went blank, and 1 s later, another trial started. There were 36 critical trials in which prime stimuli included a TNT cue and 124 filler trials in which both prime stimuli were novel words. Critical trials were divided into two types, congruent and incongruent, depending on the position of a TNT cue and probe. During congruent trials, probes were presented in the same position as TNT cues, whereas during incongruent trials, probes were presented in the opposite position of TNT cues. Cues at each level of TNT status and both probes were equally presented in the upper or lower position. Word pairs were presented once in random order; however, filler pairs were always presented during the first eight trials, and these data were excluded from later analyses.

Data Preparation

Recall latencies during the initial learning test and final recall tests were defined as the time length from cue onset to the verbal response onset. To estimate recall latencies, vocal onsets were detected manually with Audacity 2.1.1 (Audacity Team, 2015). For each sound file, the time of the vocal onset was determined manually by placing the audio track cursor at the voice sound wave onset. If a sound file contained two responses and the latter one was correct (i.e., participants corrected their responses), the onset of the second response was used for recall latency estimation. Since the recording of sound in a trial started at cue onset, the time point of vocal onset served as the recall latency of the trial.[3] To assess the effect of retrieval stopping, the recall latency of individual targets during the initial learning test was subtracted from the latency during the final recall test. The resultant values are referred to as "recall delays" in the Results section.[4] Positive recall delays indicate impaired recall, whereas negative recall delays reflect enhanced

---

[3] To evaluate the reliability of the manual estimation of recall latency, we compared the manually estimated values to the values estimated by Chronset (Roux et al., 2017), which is an automated tool for detecting speech onset. Chronset estimates an onset of the first vocal response in a sound file and returns the time length from the beginning of the file to the onset. To assess reliability, we used 2979 of 3948 (approximately 75%) responses from 47 participants (84 per participant; 36 responses in each of the initial learning and final recall tests and 12 responses in the recall test for substitutes). The other 969 responses were removed for the following reasons: responses were incorrect; responses were preceded by speech fillers; correct responses were preceded by incorrect responses in a trial (i.e., participants corrected their responses); and Chronset returned values within which participants were not able to recall targets (less than 500 ms). We conducted a linear regression analysis in which data from manual processing was an explanatory variable and data from Chronset was a dependent variable. The result ensured estimated values were closely matched between the two methods, although a few values matched poorly ($R^2 = .92$, Fig. 1).
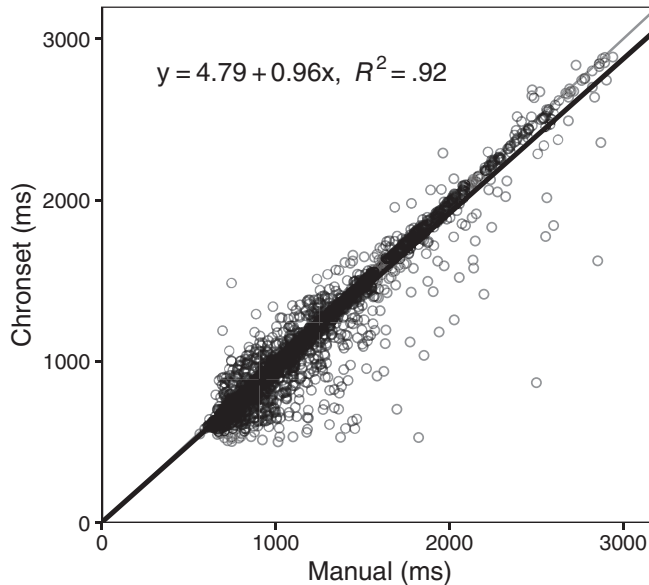
$$y = 4.79 + 0.96x, \ R^2 = .92$$

Fig. 1. Recall Latencies Estimated by Manual Processing With Audacity and by Chronset
Note. Each point represents the recall latency of an item estimated by two methods, manual processing and Chronset. Points on the thin diagonal line indicate that estimated values matched between two methods. The thick line is a linear regression line fit to the data with an equation where data estimated by manual processing is an explanatory variable and data estimated by Chronset is a dependent variable.

recall.

The data obtained during the dot-probe task were excluded based on the following criteria and procedures. First, trials were excluded in which participants could not respond to the probes correctly. Then, individual outliers were defined as RTs that deviated by more than 2.5 median absolute deviation (*MAD*) from the individual median RT and were removed (Leys et al., 2013). Eventually, 3.30 (*SD* = 2.09) critical and 6.92 (*SD* = 3.57) filler trials (approximately 9% and 6% of trials, respectively) were excluded on average.

RESULTS

Thought Evaluation Questionnaire Ratings

The patterns of scores confirmed that participants performed the TNT task as instructed, as in Experiment 1 (see Table 1). A series of 2 (No-Think strategy) × 3 (word type) ANOVAs were conducted separately for Think and No-Think item sets. For the Think condition, only a main effect of word type was significant ($F(1.28, 57.57) =$

---

[4] Although we could use recall latency estimated by Chronset for the recall delay calculation, we used the manually-estimated latency, which provided available data more than the Chronset-estimated latency (see, footnote 3). Yet, we could replace manually-estimated values with estimates by Chronset for the trials where Chronset detected vocal onsets successfully. Even if recall latency was replaced in those trials, all statistical analyses indicated the same results in terms of statistical significance. For the purpose of keeping the estimation method consistent, we computed recall delay of all trials from recall latency estimated manually.

175.31, $p < .001$, $\eta_G^2 = 0.74$). Post hoc Shaffer's tests showed that both cues and targets were rated higher than substitutes (cues vs. substitutes, $t(45) = 15.15$, $p < .001$, Hedges' $g = 3.04$; targets vs. substitutes, $t(45) = 28.43$, $p < .001$, Hedges' $g = 5.41$).

For the No-Think condition, the main effects of both factors and the interaction were significant (No-Think strategy: $F(1, 45) = 9.21$, $p = .004$, $\eta_G^2 = 0.07$; word type: $F(1.74, 78.2) = 42.12$, $p < .001$, $\eta_G^2 = 0.38$; interaction: $F(1.74, 78.2) = 27.92$, $p < .001$, $\eta_G^2 = 0.29$). Simple main effect analyses showed significant main effects of item type for two No-Think strategies (direct suppression, $F(1.33, 29.32) = 76.11$, $p < .001$, $\eta_G^2 = 0.67$; thought substitution, $F(2, 46) = 17.03$, $p < .001$, $\eta_G^2 = 0.34$). Post hoc t-tests confirmed that in the direct suppression group, cue scores were higher than target and substitute scores (cue vs. target, $t(22) = 8.76$, $p < .001$, Hedges' $g = 2.99$; cue vs. substitute, $t(22) = 9.69$, $p < .001$, Hedges' $g = 2.99$). In contrast, in the thought substitution group, cue and substitute scores were higher than target scores (cue vs. target, $t(23) = 6.50$, $p < .001$, Hedges' $g = 1.06$; substitute vs. target, $t(23) = 6.50$, $p < .001$, Hedges' $g = 2.08$).

Recall Rates and Recall Delays

As in Experiment 1, fewer No-Think targets were recalled than any other targets (Table 2). A 2 (No-Think strategy) × 3 (TNT status) ANOVA on target recall rates revealed only a significant main effect of TNT status ($F(1.17, 52.62) = 13.39$, $p < .001$, $\eta_G^2 = 0.17$). Post hoc t-tests showed that the No-Think target recall rate was significantly lower than that of any other targets (No-Think vs. Think, $t(45) = 3.99$, $p < .001$, Hedges' $g = 0.74$; No-Think vs. Baseline, $t(45) = 3.52$, $p = .002$, Hedges' $g = 0.74$).

Recall delay results also showed a strong No-Think effect (Fig. 2). Two participants' data were excluded from analyses because not all of their verbal responses were recorded due to a technical error. A 2 (No-Think strategy) × 3 (TNT status) ANOVA on recall delays showed a significant main effect of TNT status ($F(2, 86) = 94.25$, $p < .001$, $\eta_G^2 = 0.52$). Post hoc t-tests revealed that No-Think target recall was slower than any other target recall (No-Think vs. Think, $t(43) = 12.96$, $p < .001$, Hedges' $g = 2.19$; No-Think vs. Baseline, $t(43) = 10.01$, $p < .001$, Hedges' $g = 1.89$). In addition, Baseline target recall was slower than Think target recall ($t(43) = 2.49$, $p = .017$, Hedges' $g = 0.46$). There was also a significant interaction ($F(2, 86) = 5.94$, $p = .004$, $\eta_G^2 = 0.06$). Simple main effect analyses confirmed that No-Think recall was slower in the thought substitution group than in the direct suppression group ($F(1, 43) = 6.35$, $p = .016$, $\eta_G^2 = 0.13$). However, there were no differences in recall delays for Think and Baseline items between the two groups (Fs < 2, ps $> .15$).

Recall rates and recall latencies for substitutes confirmed that participants in both groups followed their assigned instructions. Another participant was excluded from the recall latency analyses because of a recording error. For the recall rate, Welch's t-test showed that the thought substitution group recalled more substitutes than the direct suppression group ($t(23.40) = 3.78$, $p < .001$, Hedges' $g = 1.11$). For recall latency, Welch's t-test showed that the thought substitution group recalled substitutes faster than the direct suppression group (direct suppression: mean RT = 1,489 ms, $SD = 233$; thought substitution: mean RT = 1,133 ms, $SD = 135$; $t(33.16) = 6.05$, $p < .001$, Hedges'
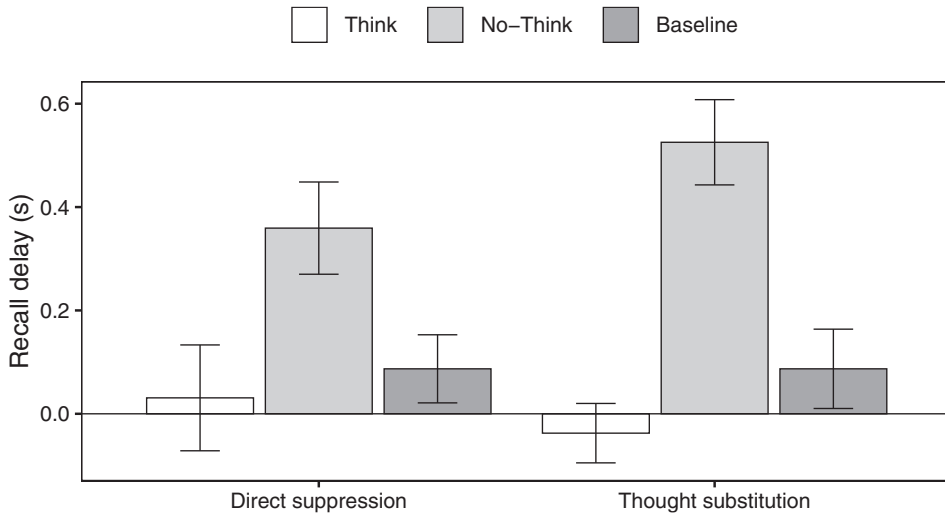
Fig. 2. Mean Recall Delays for Each Level in Experiment 2

Note. Error bars represent the 95% confidence interval of the mean.

$g = 1.79$).

RTs During the Dot-Probe Task

RTs differed significantly between the two trial types within the thought substitution group but not within the direct suppression group (Fig. 3). This tendency was confirmed by a 2 (No-Think strategy) × 3 (TNT status) × 2 (congruency) ANOVA, which showed a significant main effect of congruency ($F(1, 45) = 4.52$, $p = .039$, $\eta_G^2 = 0.01$) and a significant interaction between No-Think strategy and congruency ($F(1, 45) = 4.81$, $p = .034$, $\eta_G^2 = 0.01$). Post hoc analysis of simple main effects of the interaction confirmed that the congruency effect appeared only in the thought substitution group ($F(1, 23) = 7.69$, $p = .011$, $\eta_G^2 = 0.02$). The main effects of No-Think strategy and TNT status, and the three-way interaction were not significant (Fs < 1.1, ps > .3).

Correlations Between Performance on the Two Tasks

Correlations (skipped Pearson's correlations; Wilcox, 2016) between performance on the recall test and RTs during the dot-probe task were calculated for each group. For outlier removal, marginal medians were used. Percentile bootstrap 95% confidence intervals (CIs) were computed to estimate *p*-values (the *p*-values reported below were not adjusted). The number of bootstrap samples was 500. The recall rate data were not used in this analysis because computing marginal medians failed, likely due to a ceiling effect on the recall rates. Consistent with the inhibition transfer account, in the direct suppression group, recall delays for No-Think items were correlated positively with congruent trials RTs in which the primes contained a No-Think cue, although the correlation was not significant (*r*-skipped = .47, *p* = .18, bootstrap 95% CI [−.20, .82]). In the thought substitution group, the correlation was weaker (*r*-skipped = .23, *p* = .49,
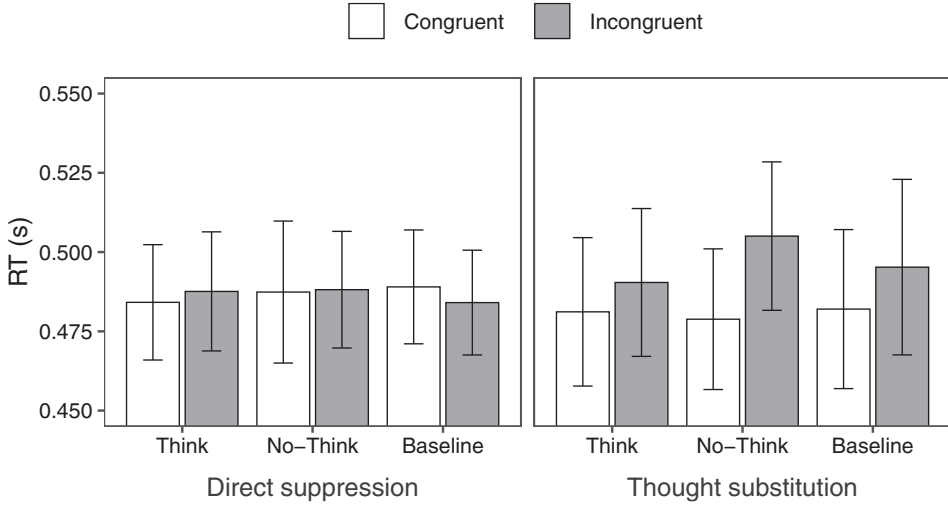
Fig. 3.   Mean RTs for Each Level During the Dot-Probe Task
Note. Error bars represent the 95% confidence interval of the mean.

Table 3.   Skipped-Correlations Between Recall Delays, Dot-Probe RTs, and Congruency Effect

	Direct suppression				Thought substitution			
	1	2	3	4	1	2	3	4
Think								
1. Recall delay	1	−.02	.08	−.13	1	.01	.09	.10
2. Congruent RT		1	.72	−.33		1	.70	−.34
3. Incongruent RT			1	.41			1	.43
4. Congruency effect				1				1
No-Think								
1. Recall delay	1	.47	.05	−.66	1	.23	.21	−.01
2. Congruent RT		1	.76	−.58		1	.85	−.18
3. Incongruent RT			1	.09			1	.36
4. Congruency effect				1				1
Baseline								
1. Recall delay	1	.00	−.30	−.32	1	.37	.32	−.03
2. Congruent RT		1	.61	−.54		1	.78	−.20
3. Incongruent RT			1	.34			1	.45
4. Congruency effect				1				1

bootstrap 95% CI [–.32, .55]). The two correlations were not significantly different (bootstrap 95% CI of correlation difference [–0.51, 0.90], p = .43). Exploratively, we computed a skipped correlation of a congruency effect (mean congruent RT – mean incongruent RT) with the recall delay for each condition. The congruency effect represents how TNT cues caught the participants' attention. The results showed a significant negative correlation between the recall delay of No-Think targets and the congruency effect for No-Think cues in the direct suppression group (r-skipped = –.66, p = .04, bootstrap 95% CI [–.82, –.05]), while the correlation was not significant in the thought substitution group (r-skipped = –.01, p = .80, bootstrap 95% CI [–.56, .51]). The two correlations were not significantly different (bootstrap 95% CI of correlation difference [–1.23, 0.16], p = .11). Table 3 shows the correlations between the recall delay, mean RTs for both congruent and incongruent trials, and the congruency effect for all conditions.

DISCUSSION

Experiment 2 was carried out to examine whether the inhibition transfer or attentional capture account was more plausible for explaining the aftereffect of retrieval stopping observed by Hertel and Hayes (2015). Retrieval stopping was followed by a dot-probe task, in which a portion of the prime stimuli contained TNT cues. First, memory impairment via retrieval stopping was successfully replicated, as in Experiment 1. Participants recalled fewer No-Think targets than Think or Baseline targets during the final recall test, although the effect sizes were smaller than those observed in Experiment 1. This is probably due to the final recall test being conducted after a dot-probe task in which participants saw all of the TNT cues. Observing these cues may have activated the corresponding targets, which aided subsequent recall. Importantly, however, the recall delay results clearly revealed memory impairment: the recall of No-Think targets was slowed by retrieval stopping compared to the recall of Baseline targets. This suggests that the effect of retrieval stopping still exists, even for correctly recalled items. In other words, an evaluation exclusively using recall rates potentially underestimates memory impairment produced by No-Think actions. Given that motor inhibition, which is involved in inhibitory mechanisms common to retrieval stopping (Depue et al., 2016; Guo et al., 2018), is assessed by RTs in the stop signal task, memory inhibition could also be assessed in the same way. This is supported by the recall delay results observed in the current study. There are some studies that did not demonstrate memory impairment in terms of recall rates (e.g., Bulevich et al., 2006). It is still possible that, in these studies, memory impairment was caused by retrieval stopping, but this effect was not strong enough to be detected by evaluating the recall rates. Although it cannot be determined whether memory impairment was reflected in recall delays in those studies, recall delay assessments appear to be an adequate complementary measure for denoting impairment.

Interestingly, the No-Think target recall delay was longer in the thought substitution group than in the direct suppression group. This indicates that thought substitution is

more effective than direct suppression, at least for the same probe test (e.g., Hertel & Calcaterra, 2005). Memory impairment by thought substitution for the same probe test is not only caused by inhibition (similar to retrieval-induced forgetting; Anderson et al., 1994; del Prete et al., 2015) but is also caused by interference from substitutes. These two factors enhanced memory impairment within the thought substitution group.

The recall delay results also provide evidence of a positive control effect (i.e., better recall of Think targets than Baseline targets). Participants recalled Think targets faster than Baseline targets. This effect was not observed for recall rates, including in Experiment 1. These recall rate results are inconsistent with previous studies. However, this does not mean that the instructions were not appropriate. The reason for a lack of positive control effect in recall rates could be that only the pairs whose targets were recalled in the initial learning test were included in the analyses to more clearly determine the retrieval stopping effect. This inclusion produced a ceiling effect, which masked the positive control effect. Despite this limitation, the recall delay results allowed us to confirm a positive control effect. In this regard, recall delays can be a complementary measure to recall rates.

For the dot-probe task, there was no direct evidence supporting either inhibition transfer or attentional capture. The inhibition transfer account predicted a two-way interaction between TNT status and No-Think strategy that was characterized by slower responses of the trials in which No-Think cues were presented as primes than in trials in which any other types of cues were presented in the direct suppression group. The attentional capture account predicted a three-way interaction that was characterized by larger congruency effect for trials where No-Think cues served as primes in the direct suppression group than any other conditions. Neither of these interactions were observed. Although the results are not compatible with both accounts, a significant two-way interaction between No-Think strategy and congruency favours the inhibition transfer account as an explanation for the aftereffect reported by Hertel and Hayes (2015) over the attentional capture account for three reasons. First, participants in the direct suppression group showed no congruency effect in the No-Think condition, whereas participants in the thought substitution group exhibited a congruency effect. This indicates that No-Think cues for direct suppression do not draw attention, whereas No-Think cues for thought substitution draw attention. This is completely contrary to the attentional capture account, which assumes that the aftereffect (delayed rating) was caused by excessive attention on cues to achieve direct suppression.

Second, the two groups showed different congruency effects for trials in which Think and Baseline cues were primes: congruency effects were observed in the thought substitution group but not in the direct suppression group. Treatments for Think and Baseline cues during the TNT task should have been identical between the two groups, and thus similar congruency effects should have been observed. The lack of a congruency effect for the two types of trials in the direct suppression group might be due to an inhibition transfer that was mistakenly elicited by Think and Baseline cues. The prime stimulus duration (500 ms) was not enough to discriminate which type of cue was presented; thus, participants were only able to discern whether or not studied items were

presented. Therefore, whenever a TNT cue was presented, participants exerted inhibitory mechanisms that were learned during the preceding TNT phase. This resulted in the lack of a congruency effect, in any condition, within the direct suppression group. The speculative nature of this interpretation requires further study.

Finally, the results of the correlation analyses between recall and dot-probe task performance favour the inhibition transfer account. For the direct suppression group, a negative correlation between recall delay and congruency effect was found, which indicates that individuals who suppressed targets more successfully had weaker or inverse congruency effects (i.e., longer mean congruent RTs compared to mean incongruent RTs). In contrast, this relationship was not observed in the thought substitution group that demonstrated greater memory impairment compared to the direct suppression group, although the correlations were not significantly different between the two groups. These results are not consistent with the attentional capture account predicting that cues for well-trained No-Think attempts draw one's attention so that the recall delay is positively correlated with the congruency effect. Instead, the results partly support the inhibition transfer account. Weaker or inverse congruency effects are derived from shorter incongruent RTs and/or longer congruent RTs. Although it was not significant, a moderate positive correlation between recall delays and congruent RTs in the No-Think condition implies that longer RTs in congruent trials might be derived from precedent presentations of cues for well-trained No-Think attempts. Responses to probes at the same location as No-Think cues might be prevented by inhibitory control triggered by No-Think cues via their association that was formed through repeated No-Think attempts during the TNT task that involve inhibitory control (e.g., Depue et al., 2016). Even though the correlation results are not direct evidence of inhibition transfer, they are not compatible with the attentional capture account at all.

Taken together, although the results of the dot-probe task were not fully consistent with the predictions, the results favour the inhibition transfer account more than the attentional capture account. In conjunction with the results of Hertel and Hayes (2015), it is reasonable to conclude that inhibition transfer may underlie the direct suppression aftereffect.

GENERAL DISCUSSION

The aim of the current study was to determine whether the inhibition transfer account or attentional capture account is more plausible for explaining the aftereffect of direct suppression observed in Hertel and Hayes (2015). Two experiments were conducted: Experiment 1 was designed to replicate basic TNT findings, and Experiment 2 directly examined the research question. First, the current study replicated the memory impairment effect via retrieval stopping in both experiments. In addition, Experiment 2 revealed that recall delay (or latency) analyses are appropriate for indexing memory inhibition. As shown in previous studies, repeated retrieval stopping causes forgetting regardless of the stopping strategies used (e.g., Anderson & Huddleston, 2012), although

different inhibitory mechanisms seem to underlie forgetting by direct suppression and thought substitution (Benoit & Anderson, 2012). During direct suppression, inhibitory control over retrieval is exerted in response to the counter-intentional retrieval of targeted memories (Benoit et al., 2015; Levy & Anderson, 2012), leading to overall suppression of the memory system that prevents not only retrieval but also encoding (Hulbert et al., 2016). The interruption of reconstruction or rehearsal loops by systemic suppression presumably impairs traces of the targeted memories, which results in the memories being forgotten. In contrast, forgetting by thought substitution is based on competition-based inhibitory processes (Anderson, 2003; Benoit & Anderson, 2012; del Prete et al., 2015). Target items are so strongly associated with cues and easy to retrieve that they interfere with the retrieval of substitutes that have relatively weak associations with cues. To resolve the interference from targets and achieve substitute retrieval, memory traces of targets are inhibited, resulting in the memories being forgotten (but see Raaijmakers & Jakab, 2013). It should be noted that the present results cannot conclude that forgetting was caused by impaired memory traces. Instead, it is possible that forgetting was derived from other factors, such as unlearned associations between cues and targets (and inhibition transfer, as discussed below), or their combinations because memory performance was evaluated only in the recall test with the same cues as used during the TNT task.

Regarding the research question, the current study provides supporting evidence that the aftereffect observed in Hertel and Hayes (2015) might be caused by inhibition transfer, which suggests that inhibitory control on retrieval processes during No-Think attempts is also exerted on task performance mostly irrelevant to retrieval (i.e., valence rating), leading to impaired performance (i.e., slower response). Inhibition transfer might occur through No-Think cues. The presentation of No-Think cues was accompanied by inhibitory control on retrieval repeatedly during the TNT task, creating strong associations between No-Think cues and inhibitory control. In the following task, through the associations, the presentation of the cues triggered an inhibitory system, which exerted control over processes for the task. Importantly, inhibition transfer is likely to occur outside of one's intention. In the flanker and dot-probe tasks subsequent to the TNT task, participants were not given any instruction that would encourage them to exert inhibitory control. Previous studies have shown that control processes are exerted unintentionally by associated primes such as attentional orientation (Jiang et al., 2006), task set preparation (Lau & Passingham, 2007; Weibel et al., 2013), task switching (Reuss et al., 2011), error detection (Charles et al., 2013), conflict adaptation (van Gaal, Lamme, & Ridderinkhof, 2010), and even stop signal inhibition (van Gaal et al., 2008; van Gaal, Ridderinkhof, et al., 2010). Recently, Salvador et al. (2018) observed that control processes on memory are also automatically elicited when primes associated with a No-Think instruction appeared subliminally. In that study, participants were instructed to retrieve or stop retrieving targets from cues depending on shapes presented at the beginning of the trials (a square for Think cues and a diamond for No-Think cues, or vice versa). These prime shapes were presented in a visible way for some targets, but they were masked for the other targets. In the subliminal trials, participants performed a grammatical gender determination task for cue words. During the final recall test, participants recalled fewer

targets in the condition where an invisible No-Think prime was presented. Although in Hertel and Hayes (2015) and the current study primes for inhibitory control (i.e., No-Think cues) were presented in a visible way, the fact that participants were not given control instructions ensures that performance was unintentionally affected by TNT cues. Rather, the findings in Hertel and Hayes (2015) and the current study supporting inhibition transfer imply that inhibitory control is exerted automatically, even when primes associated with control are visible.

The characteristic of inhibition transfer triggered by No-Think cues suggests that inhibition transfer underlies forgetting in the same cue test as well as the mechanisms proposed by previous studies, such as unlearned associations between cues and targets and impaired memory traces of targets (Anderson & Huddleston, 2012). Namely, the presentation of No-Think cues during the same cue test drives inhibitory control over the retrieval processes as in No-Think trials of the preceding TNT task, which hinders individuals from recalling the corresponding targets even if individuals were encouraged to recall it. Therefore, another test method in which TNT cues were not presented, such as the independent probe test (Anderson & Spellman, 1995), should be employed to examine whether retrieval stopping impairs target memory traces.

The inhibition transfer account assumes that well-trained control over a cognitive process is exerted on a process with little relevance; this supports the concept of the domain-general nature of inhibitory control. The idea of domain generality has been supported by evidence of neural correlates involved with control processes overlapping across retrieval stopping, motor inhibition, and emotion regulation (Banich & Depue, 2015; Depue et al., 2016). While the previous findings support domain generality using neuroimaging and correlational methods, inhibition transfer has behavioural support, presumably highlighting the causality of cross-domain influences of inhibition. In this regard, inhibition transfer is an important phenomenon whereby underlying mechanisms should be investigated further.

Although the present findings favour an inhibition transfer account, namely, from the overall difference in congruency observed between the two groups, it is also possible that the current results are attributable to a task-set priming effect of distinct No-Think strategies on the dot-probe task. For direct suppression, participants are instructed to focus on only one representation (i.e., cue) and discard others. In contrast, for thought substitution, participants are instructed to actively construct cue-substitute associations. The different instructions created different task sets for the following dot-probe task; the direct suppression group was encouraged to concentrate on judging upcoming probes and ignore primes, while the thought substitution group was encouraged to process all stimuli, including primes. The differential task sets result in the lack of a congruency effect in the direct suppression group and a congruency effect in the thought substitution group. However, the task set account for overall differences in the congruency effect between groups is incompatible with the absence of an overall difference between groups in Hertel and Hayes (2015). This might be due to the difference in TNT cue presentation methods in the task subsequent to the TNT task. TNT cues were presented briefly (500 ms) prior to probe stimuli in the present study, whereas the cues were presented as flankers along

with words to be rated until ratings for the words were finished (for 2,000 ms on average) in Hertel and Hayes (2015). Compared to cues presented briefly and separately from stimuli to be responded, cues presented for longer times along with to-be-responded stimuli were more difficult to ignore. For this reason, all cues presented as flankers were processed, and some characteristics of No-Think cues created through the TNT task influenced task performance (i.e., rating). If this task-set priming is genuine, the dot-probe task might not have been appropriate for investigating the underlying direct suppression aftereffect mechanism because the primes were not processed. The possibility that differences in instructions could influence the task set is worth examining in future studies.

Alternatively, the lack of a congruency effect in the dot-probe task in the direct suppression group is derived from ignoring items based on a template for rejection that was created through the TNT task. Recent attention studies have shown that one can ignore features/items intentionally by acquiring a template for rejection through repeated attempts (e.g., Cunningham & Egeth, 2016). Even though participants were instructed to focus on presented cues in the TNT trials, it is possible that participants perform retrieval stopping successfully by ignoring the cues during direct suppression. In the repeated No-Think attempts, participants created a template for rejection, which encouraged participants to ignore the No-Think cues in the dot-probe task, resulting in the absence of a congruency effect. A template for rejection of No-Think cues can also account for poorer recall of suppressed targets in the same cue test. Although cues should be attended to recall the associated targets, the template for rejection prevents participants from attending and processing No-Think cues, leading to delayed recall or forgetting. Unlike the current results, it is not consistent with the results of the flanker task reported by Hertel and Hayes (2015). If a template for rejection of No-Think cues is the case, words flanked by No-Think cues should be rated faster than those flanked by Think or Baseline cues because participants would have ignored the No-Think cues and would not have been distracted by them. Nevertheless, it is worthwhile to examine whether a template for rejection is created through repeated direct suppression. It might reveal the mechanisms of retrieval stopping, subsequent forgetting in the same cue test, and the consequences of retrieval stopping in the long run.

In addition to the limitation of the current results that other mechanisms can account for, there might be a sensitivity issue regarding the dot-probe task. As shown in Fig. 3, the congruency effect (incongruent RT – congruent RT) was approximately 25 ms, although trial-by-trial variance within each participant was larger than the effect (mean trial-by-trial SD was 74 ms). Furthermore, there were large individual differences in mean RTs. Although the large variance contributed to the significant correlation of the congruency effect with recall delay, it might have attenuated the congruency effect. Hedge et al. (2018) demonstrated that robust cognitive tasks usually lead to low individual differences, in which it is difficult to make a correlation-based prediction simultaneously. The current results of the dot-probe task exhibited the opposite pattern, which might make it difficult to find another factor (i.e., TNT status and No-Think strategy here) to interact with the congruency factor from the series of ANOVAs.

In summary, the current study provided evidence supporting the inhibition transfer account regarding the aftereffect of direct suppression observed by Hertel and Hayes (2015). Inhibition transfer is, on one hand, considered a cost of retrieval stopping because cues associated with stopping unexpectedly impair task performance in another context. On the other hand, inhibition transfer might play a key role in behaviourally demonstrating the domain generality of inhibition during retrieval stopping.

AUTHOR'S CONTRIBUTION

Both authors contributed to the conception, experimental design, interpretation of data, and manuscript preparation. S.N. took primary responsibility for data collection and analysis.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

Amano, S., & Kondo, T. (Eds.). (1999). *Nihongo-no goitokusei: Tango shinmitsudo* [Lexical properties of Japanese: Word familiarity] (Vol. 1). Sanseido.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, *49*(4), 415–445. https://doi.org/10.1016/j.jml.2003.08.006

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1063–1087. https://doi.org/10.1037/0278-7393.20.5.1063

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, *410*(6826), 366–369. https://doi.org/10.1038/35066572

Anderson, M. C., & Huddleston, E. (2012). Towards a cognitive and neurobiological model of motivated forgetting. In R. F. Belli (Ed.), *True and false recovered memories* (pp. 53–120). Springer. https://doi.org/10.1007/978-1-4614-1195-6_3

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*(1), 68–100. https://doi.org/10.1037/0033-295X.102.1.68

Audacity Team. (2015). *Audacity (R): Free audio editor and recorder* (Version 2.1.1) [Computer application]. Audacity. https://audacityteam.org/

Banich, M. T., & Depue, B. E. (2015). Recent advances in understanding neural systems that support inhibitory control. *Current Opinion in Behavioral Sciences*, *1*, 17–22. https://doi.org/10.1016/j.cobeha.2014.07.006

Benoit, R. G., & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron*, *76*(2), 450–460. https://doi.org/10.1016/j.neuron.2012.07.025

Benoit, R. G., Hulbert, J. C., Huddleston, E., & Anderson, M. C. (2015). Adaptive top–down suppression of hippocampal activity and the purging of intrusive memories from consciousness. *Journal of Cognitive Neuroscience*, *27*(1), 96–111. https://doi.org/10.1162/jocn_a_00696

Bulevich, J. B., Roediger, H. L., Balota, D. A., & Butler, A. C. (2006). Failures to find suppression of

episodic memories in the think/no-think paradigm. *Memory & Cognition*, *34*(8), 1569–1577. https://doi.org/10.3758/BF03195920

Charles, L., van Opstal, F., Marti, S., & Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *NeuroImage*, *73*, 80–94. https://doi.org/10.1016/j.neuroimage.2013.01.054

Cunningham, C. A., & Egeth, H. E. (2016). Taming the white bear: Initial costs and eventual benefits of distractor inhibition. *Psychological Science*, *27*(4), 476–485. https://doi.org/10.1177/0956797615626564

del Prete, F., Hanczakowski, M., Bajo, M. T., & Mazzoni, G. (2015). Inhibitory effects of thought substitution in the think/no-think task: Evidence from independent cues. *Memory*, *23*(4), 507–517. https://doi.org/10.1080/09658211.2014.907429

Depue, B. E., Orr, J. M., Smolker, H. R., Naaz, F., & Banich, M. T. (2016). The organization of right prefrontal networks reveals common mechanisms of inhibitory regulation across cognitive, emotional, and motor processes. *Cerebral Cortex*, *26*(4), 1634–1646. https://doi.org/10.1093/cercor/bhu324

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Guo, Y., Schmitz, T. W., Mur, M., Ferreira, C. S., & Anderson, M. C. (2018). A supramodal role of the basal ganglia in memory and motor inhibition: Meta-analytic evidence. *Neuropsychologia*, *108*, 117–134. https://doi.org/10.1016/j.neuropsychologia.2017.11.033

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review*, *12*(3), 484–489. https://doi.org/10.3758/BF03193792

Hertel, P. T., & Hayes, J. A. (2015). Distracted by cues for suppressed memories. *Psychological Science*, *26*(6), 775–783. https://doi.org/10.1177/0956797615570711

Hulbert, J. C., Henson, R. N., & Anderson, M. C. (2016). Inducing amnesia through systemic suppression. *Nature Communications*, *7*, Article 11003. https://doi.org/10.1038/ncomms11003

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(45), 17048–17052. https://doi.org/10.1073/pnas.0605678103

Kassambara, A. (2020). *rstatix: Pipe-friendly framework for basic statistical tests*. Retrieved from https://CRAN.R-project.org/package=rstatix

Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *Journal of Neuroscience*, *27*(21), 5805–5811. https://doi.org/10.1523/JNEUROSCI.4335-06.2007

Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *Journal of Neuroscience*, *32*(47), 16785–16794. https://doi.org/10.1523/JNEUROSCI.2640-12.2012

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15–20. https://doi.org/10.1037/0021-843X.95.1.15

Nørby, S. (2015). Why forget? On the adaptive value of memory loss. *Perspectives on Psychological Science*, *10*(5), 551–578. https://doi.org/10.1177/1745691615596787

Nørby, S. (2018). Forgetting and emotion regulation in mental health, anxiety and depression. *Memory*, *26*(3), 342–363. https://doi.org/10.1080/09658211.2017.1346130

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, Article 10. https://doi.org/10.3389/neuro.11.010.2008

R Core Team. (2020). *R: A language and environment for statistical computing*. Retrieved from https://www.R-project.org/

Raaijmakers, J. G., & Jakab, E. (2013). Rethinking inhibition theory: On the problematic status of the

inhibition theory for forgetting. *Journal of Memory and Language*, *68*(2), 98–122. https://doi.org/10.1016/j.jml.2012.10.002

Reuss, H., Kiesel, A., Kunde, W., & Hommel, B. (2011). Unconscious activation of task sets. *Consciousness and Cognition*, *20*(3), 556–567. https://doi.org/10.1016/j.concog.2011.02.014

Roux, F., Armstrong, B. C., & Carreiras, M. (2017). Chronset: An automated tool for detecting speech onset. *Behavior Research Methods*, *49*(5), 1864–1881. https://doi.org/10.3758/s13428-016-0830-1

Salvador, A., Berkovitch, L., Vinckier, F., Cohen, L., Naccache, L., Dehaene, S., & Gaillard, R. (2018). Unconscious memory suppression. *Cognition*, *180*, 191–199. https://doi.org/10.1016/j.cognition.2018.06.023

van Gaal, S., Lamme, V. A., & Ridderinkhof, K. R. (2010). Unconsciously triggered conflict adaptation. *PLOS ONE*, *5*(7), Article e11508. https://doi.org/10.1371/journal.pone.0011508

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *Journal of Neuroscience*, *28*(32), 8053–8062. https://doi.org/10.1523/JNEUROSCI.1278-08.2008

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. (2010). Unconscious activation of the prefrontal no-go network. *Journal of Neuroscience*, *30*(11), 4143–4150. https://doi.org/10.1523/JNEUROSCI.2992-09.2010

Wang, Y., Cao, Z., Zhu, Z., Cai, H., & Wu, Y. (2015). Cue-independent forgetting by intentional suppression – Evidence for inhibition as the mechanism of intentional forgetting. *Cognition*, *143*, 31–35. https://doi.org/10.1016/j.cognition.2015.05.025

Weibel, S., Giersch, A., Dehaene, S., & Huron, C. (2013). Unconscious task set priming with phonological and semantic tasks. *Consciousness and Cognition*, *22*(2), 517–527. https://doi.org/10.1016/j.concog.2013.02.010

Wilcox, R. R. (2016). *Introduction to robust estimation and hypothesis testing* (4th ed.). Academic Press.