



BIROn - Birkbeck Institutional Research Online

Zivony, Alon and Eimer, Martin (2020) Perceptual competition between targets and distractors determines working memory access and produces intrusion errors in RSVP tasks. *Journal of Experimental Psychology: Human Perception and Performance* , ISSN 0096-1523. (In Press)

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/32734/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html> or alternatively contact lib-eprints@bbk.ac.uk.

**Perceptual competition between targets and distractors determines working memory access
and produces intrusion errors in RSVP tasks**

Alon Zivony* & Martin Eimer

Department of Psychological Sciences, Birkbeck College, University of London,
Malet Street, London WC1E 7HX, UK

*This is the accepted authors' manuscript, not the final published article. The published article will
be available at:*

Journal of Experimental Psychology: Human Perception and Performance

Word count: 14,343

* Corresponding author

Phone: 0044 20 76316522

Fax: 0044 20 76316312

Email: alonzivony@gmail.com

Abstract

When a target and a distractor that share the same response dimension appear in rapid succession, participants often erroneously report the distractor instead of the target. Using behavioral and electrophysiological measures, we examined whether these intrusion errors occur because the target is often not encoded in working memory (WM) or are generated at later post-encoding stages. In four experiments, participants either provided two guesses about the target's identity, or had to select the target among items that did not include the potential intruder. Results showed that the target did not gain access to WM on a substantial number of trials where the distractor was encoded. This was also confirmed with an electrophysiological marker of WM storage (CDA component). These findings are inconsistent with post-encoding accounts of distractor intrusions, which postulate that competitive interactions within WM impair awareness of the target, the precision of target representations, or result in the target being dropped from WM. They show instead that target-distractor competition already operates at earlier perceptual stages, and reduces the likelihood that the target gains access to WM. We provide a theoretical framework to explain these findings and how they challenge contemporary models of temporal attention.

Keywords: distractor intrusions, RSVP, temporal selection, working memory

Public Significance Statement

When a target and distractors are presented in rapid succession at the same location, participants often mistakenly report one of these distractors as being the target. These distractor intrusions reflect a robust limitation of attentional control in the time domain. Our study suggests that distractor intrusions are caused by competitive interactions during perceptual processing that can block the target's access to working memory.

Goal-directed behavior often requires task-relevant information to be differentiated from irrelevant information. In many visual tasks, identifying a target object necessitates its segregation from the background and from surrounding distractor objects. Visual attention promotes this process by selectively biasing processing in favour of targets relative to distractors. When the target is uniquely defined by a basic visual feature, such as colour, orientation, or shape, attention will be automatically guided to its location (Wolfe & Horowitz, 2004; 2017). Selection by a target-defining feature is similarly effective in rapid serial visual presentation (RSVP) tasks where multiple objects appear successively and rapidly at the same location, and attention needs to be guided to the right object and the right moment in time (e.g., Folk et al., 2002).

The allocation of attention to target objects facilitates the processing of these objects, at the expense of other currently-task irrelevant distractors that can be safely ignored. However, some distractors are more difficult to ignore than others. In attentional selection tasks where multiple stimuli are presented simultaneously, distractors that share the target category can impede target identification (e.g., Broadbent, 1982). For example, when participants have to identify a target digit, spatially adjacent distractor digits produce stronger interference with performance relative to distractor letters (Jonides & Gleitman, 1972), even when the location of the target is fixed and known in advance (Avital-Cohen & Tsal, 2016; Chanceaux et al, 2014). Pronounced interference effects from category matching distractors also emerge in RSVP tasks. When participants have to report the identity of a target digit that is presented among distractor digits in an RSVP stream, they will often erroneously report the identity of temporally adjacent distractors. Such *distractor intrusions* have been reported in multiple studies (e.g., Botella & Eriksen, 1992; Botella et al., 2001; Chun, 1997; Intraub, 1985; Gathercole & Broadbent, 1984; Goodbourn et al., 2016; Kikuchi, 1996; Popple & Levi, 2007; Vul et al., 2009).

In a recent study (Zivony & Eimer, in press), we employed tasks with two lateral RSVP streams where participants had to report the identity of a target digit (defined by its surrounding shape, see Figure 1, or by its colour) in one of these streams. Relative to trials where the target was followed by a neutral distractor that did not match the target category (a letter), accuracy was strongly reduced on trials where the target was followed by a potentially intruding distractor (another digit). The vast majority of errors on these trials were reports of this post-target distractor (PTD). The prevalence of such PTD intrusion errors poses a serious challenge to the assumption shared by many models of temporal attention that a single feature-defined target can be easily distinguished from surrounding distractors, even at high presentation rates (Chun & Potter, 1995; Di Lollo et al., 2005; Jolicœur & Dell’Acqua, 1998; Taatgen et al., 2009).

However, this PTD intrusion effect can be accommodated by models that consider the time course of attentional processes that are triggered once a target-defining selection feature is detected (Olivers & Meeter, 2008; Reeves & Sperling, 1986; Shih, 2008; Shih & Sperling, 2002; Sperling & Weichselgartner, 1995; Wyble et al., 2011). These models postulate the existence of *attentional episodes* during which visual activity is amplified. This amplification rapidly builds up following the detection of a selection feature and then gradually dissipates over approximately 200 ms. Representations of visual items that appear inside the attentional episode (regardless of their identity) are enhanced, making them more likely to be encoded in working memory (WM). In contrast, visual representations of items that appear outside of the attentional episode remain weak, unstable, and transient, are affected by visual masking, and are therefore unlikely to gain access to WM. Thus, when a target is immediately followed by a category-matching PTD, both items are processed within the same attentional episode, and both are therefore strong candidates for encoding into WM and for controlling subsequent response selection stages. We have previously

suggested that correct responses occur on trials where the target entered WM, whereas PTD intrusion errors occur when the PTD is encoded instead (Zivony & Eimer, in press). Furthermore, we proposed that this depends on the speed with which attention is engaged at the target location and an attentional episode is triggered. Because the target and the PTD are represented in the same retinotopic coordinates in the visual cortex, they compete with each other (Wyble et al., 2009, 2011). On trials where an attentional episode is triggered early, target representations will be more strongly facilitated than representations of the PTD, thus resolving their competition in favour of the target. In contrast, when attentional engagement is delayed, the PTD representation will be more strongly activated, and is therefore more likely to gain access to WM.

To test this hypothesis, we measured the onset latency of N2pc components triggered by target frames in the RSVP streams as an event-related potential (ERP) marker of attentional engagement speed. The N2pc is an electrophysiological marker of the allocation of attention to objects with target-defining attributes (Eimer, 1996; Eimer et al., 2009; Woodman & Luck, 1999), which has recently also been employed to measure the engagement of attention (Zivony et al., 2018). In three experiments, we measured N2pc components separately for trials with correct responses and trials with PTD intrusion errors, and found systematic N2pc onset latency differences, with consistently earlier N2pc onsets on correct as compared to intrusion trials (Zivony & Eimer, in press). These results support our hypothesis that the presence versus absence of PTD intrusion errors is linked to trial-by-trial temporal variability in the onset of attentional episodes, by demonstrating that this type of variability can bias the competition between targets and PTDs in favour of one of these items.

However, what remains unclear is at what stage of processing this competition takes place. We proposed a model where these items compete at a relatively early stage, prior to WM encoding,

and the outcome of this competition determines whether the target or the PTD will gain entry into WM (Zivony & Eimer, in press). This *encoding competition* hypothesis is consistent with our N2pc results, but it is not the only viable account. An alternative possibility is that both items enter WM, and that the competitive interactions between them that determine which item will be reported only occur after they have been encoded in WM. This *post-encoding competition* hypothesis is supported by the common pattern of results in two-target RSVP paradigms where the temporal proximity of the two targets is manipulated. When these targets are separated by more than 500 ms, accuracy in reporting both of them is high. When this interval is shortened, accuracy drops dramatically for the second target (attentional blink, e.g., Raymond et al., 1992). Importantly, when both targets are presented in immediate succession within about 100 ms, accuracy for the second target recovers (lag-1 sparing, e.g., Visser et al., 1999). This effect suggests that when two objects appear within the same attentional episode, they can be both encoded in WM. The observation that in such lag-1 sparing situations, the reported order of the two targets is often reversed (Hommel & Akyürek, 2005) could reflect a competitive bias in WM for the representation of the second target, resulting in its prior entry into visual awareness and response selection (see Hilkenmeier et al., 2012, for a similar suggestion). An analogous prior entry effect could be responsible for PTD intrusion errors, when the PTD is mistakenly perceived to appear first and therefore misidentified as the target. Another competitive effect observed in lag-1 sparing situations is that reported target visibility and confidence are reduced relative to situations when both targets are separated by at least 500 ms (Pincham et al., 2016; Recht et al., 2019). According to Pincham et al. (2016), participants may often consciously perceive only one of the two targets even when both are stored in WM. Such an ‘experiential blink’ within WM could also account for PTD intrusions, when only the PTD is available for conscious report.

Thus, it remains unclear whether PTD intrusion errors occur when the target fails to gain access to WM, or because of subsequent competitive interactions between target and PTD representations within WM. The goal of the present study was to distinguish between these alternative encoding and post-encoding competition accounts, which is important for two reasons. First, given the robustness of the PTD intrusion effect, it can potentially open new avenues of research into the temporal selectivity of attentional processes. However, this requires that the mechanisms that underlie this effect are fully understood. Second, given the similarity of distractor intrusion effects and phenomena such as lag-1 sparing that are observed in two-target RSVP tasks, investigating whether similar or different processes operate in these cases can increase the scope and the generalizability of current models of temporal attention (Olivers & Meeter, 2008; Wyble et al., 2011).

The key difference between the encoding and post-encoding competition accounts concerns the number of items that are encoded in WM on trials where a target is followed by a category-matching PTD. According to the encoding competition account, only one of these items will be encoded and thus become available for report. The post-encoding account assumes that the presence of a PTD does not prevent the target from gaining access to WM, and vice versa. More specifically, these two accounts disagree on whether the target item is represented in WM on trials where PTD intrusion errors occur. In the present study, we used novel experimental procedures designed to test this and other related questions.

Experiment 1

In our previous study (Zivony & Eimer, in press), participants always reported only a single item on each trial. Even though PTD intrusion errors occurred on a large number of trials, the

target may have still been encoded into WM and have remained potentially available for report on these trials. Since only a single report was required, participants may have relied on additional criteria to decide which of the two encoded items was the target on any given trial. To test this possibility, we used the same general design as in our earlier study, but introduced a change to the report procedures (see Figure 1). Even though participants were instructed to detect and report a single target digit on each trial, they were now asked to provide two different guesses (instead of one target report) at the end of each trial, in order to maximize their chances to report the target. This allowed us to measure the probability of reporting the PTD digit as a second response following correct reports (intrusion | correct 1st response), and the probability of reporting the target as a second response on trials where the PTD digit was reported first (correct | intrusion 1st response). These conditional probabilities quantify the likelihood that one of two successive digits was encoded on trials where the other digit was also encoded, which is central to differentiating between the encoding competition and post-encoding competition accounts. According to the post-encoding competition account, these two items do not compete for encoding, and should therefore both gain access to WM. As other factors also affect target identification, the conditional probabilities on trials with a PTD digit are likely to remain well below 100%. However, and critically, the post-encoding account predicts that the probabilities on these trials should match the probability of correct first responses on baseline trials, as targets and PTDs do not compete for access to WM on both types of trials. On the other hand, the encoding competition account predicts that only one item (either the target or the PTD) is encoded on any given trial. In this case, the conditional probabilities of reporting either of these items when the other one was reported first should not exceed chance performance.

A similar two-response procedure was previously used by Vul et al. (2009) in order to test

whether participants were more likely to choose items for their second report that were temporally adjacent to the item that was reported first. No such temporal associations between both reports were found. Although these authors interpreted their results as evidence for the encoding of multiple items within a temporally fixed attentional window, the RSVP streams in their study contained only distractors that shared the target's response dimension. For this reason, they could not measure the accuracy of reporting a specific target in the absence versus presence of other potentially intruding items, and assess how reports are affected by the competition between the target and temporally adjacent distractors.

Method

Sample size selection

To calculate the sample size required, we focused on the two comparisons that are most relevant for distinguishing the encoding and post-encoding accounts, which are the comparisons of the conditional accuracy of second responses on trials with a potentially intruding PTD to (i) the accuracy of these responses on baseline trials and to (ii) chance accuracy (20%). Using the data from Zivony & Eimer (in press, Experiment 1A), baseline accuracy was estimated to be 77%. For comparison (i), error variance was estimated based on the comparison of accuracy rates between digit distractor and letter distractor conditions ($Sd = 12.5\%$). For comparison (ii), error variance was estimated based on the variance in accuracy on the digit distractor condition ($S = 18.6\%$). Because the post-encoding account predicts a small or no difference for comparison (i), whereas the encoding account predicts small or no difference for comparison (ii), we opted for the weakest possible scenario for obtaining reliable differences where the actual accuracy on trials with a potential intruder was exactly at the mid-point between baseline and chance accuracy (i.e., 48.5%).

Based on these data, we calculated the sample size required to observe significant effects for both comparisons using G*Power (Faul et al., 2013), with an alpha of .05 and power of .80. The minimum sample size was 4 for comparison (i) and 6 for comparison (ii). Since there were substantial differences between our previous experiment and the current Experiment 1 (most notable, one versus two response options), we opted for a substantially larger sample size of 16 participants.

Participants

Participants were 16 (11 women) volunteers ($M_{age} = 24.4$, $SD = 5.25$) who participated for a payment of £5. All reported normal or corrected-to-normal visual acuity. All methods used in this experiment, and subsequent experiments, were approved by the institution's departmental ethical guidelines committee at Birkbeck, University of London.

Apparatus

Stimuli were presented on a 24-inch BenQ monitor (100 Hz; 1920×1080 screen resolution) attached to a SilverStone PC, with participant viewing distance at approximately 80 cm. Manual responses were registered via a standard computer keyboard.

Stimuli and design

The sequence of events is illustrated in Figure 1A. Each trial began with the presentation of a fixation display (a grey $0.75^\circ \times 0.75^\circ$ "+" sign at the center of the screen). After 500 ms, two lateral RSVP streams including 8 to 11 frames appeared along with the fixation cross. Frames consisted of two alphanumeric characters (1° in height) appearing at a center-to-center distance of 3.5° to the

left and right of fixation. Each frame appeared for 50 ms, followed by an ISI of 50 ms. All stimuli in the RSVP streams were grey (CIE colour coordinates: 0.309/.332, luminance 46.6 cd/m²).

On each trial, a digit target was presented unpredictably in one of two RSVP streams on the left or right side. This target appeared inside a pre-specified shape (circle or square; selection feature), and participants had to report its numerical value (response feature) by pressing the corresponding keyboard button. For half of all participants, the target-defining selection feature was the square, and for the other half, the selection feature was the circle. Participants were instructed that their goal was to detect and report the single target object, but were asked to provide two unique responses in their attempt to do so. They were told that correct identification on the first response was preferable, but that they should use their second guess to maximize their chances of reporting the target, as they might not be sure of its identity on some trials. These two responses were executed without time pressure at the end of each trial. The first response was prompted by a response screen that contained all six possible digits in a row, 2.5° above fixation, with a center-to-center distance between each digit of 1.6° (Figure 1C). Once the first response choice was made, the chosen digit was crossed out, and participants then had to choose a second option from the remaining five digits. Following the second response, a blank screen appeared for 500 ms, after which a new trial began.

The experiment consisted of 20 practice trials and 300 experimental trials divided to 50-trial blocks. Participants were allowed to take self-paced breaks between blocks. Digits (including the target and post-target digit distractor, if present) were drawn without replacement from a limited set of six digits (2, 3, 4, 6, 7, and 8). Letters were randomly selected without replacement from a 23-letter set (all English alphabet letters, excluding I, X, and O). The target digit appeared with equal probability and unpredictably in the 5th, 6th, 7th, or 8th frame, either in the left or right RSVP

stream. This target frame contained one digit and one letter, which appeared within two different outline shapes (square: 1.2° in side, and circle: 1.3° in diameter, line width for both: 4 pixel). The digit was always presented within the pre-specified target shape, and the letter within the other shape. The frame immediately preceding the target frame always included two letters (to prevent any pre-target intrusion errors). All other pre-target frames were equally likely to contain two letters, or one digit and one letter (with digit and letter location randomly selected for each frame). The target frame was always followed by three additional frames. On two thirds of all trials, the frame immediately following the target contained a digit (i.e., a potentially intruding PTD) in the same location as the preceding target digit, so that PTD intrusion errors were possible (Figure 1A). On the remaining randomly intermixed one third of trials, this frame contained two letters (i.e., the PTD was category-nonmatching, Figure 1B). The two final frames on each trial always included two letters.

Participants were informed that target digits were equally likely to appear in the left or right RSVP stream, and that task-irrelevant digits would appear prior to the target. This ensured that attentional allocation processes would be guided by the selection feature (circle or square), rather than by alphanumeric category (i.e., attending to the first digit in the stream).

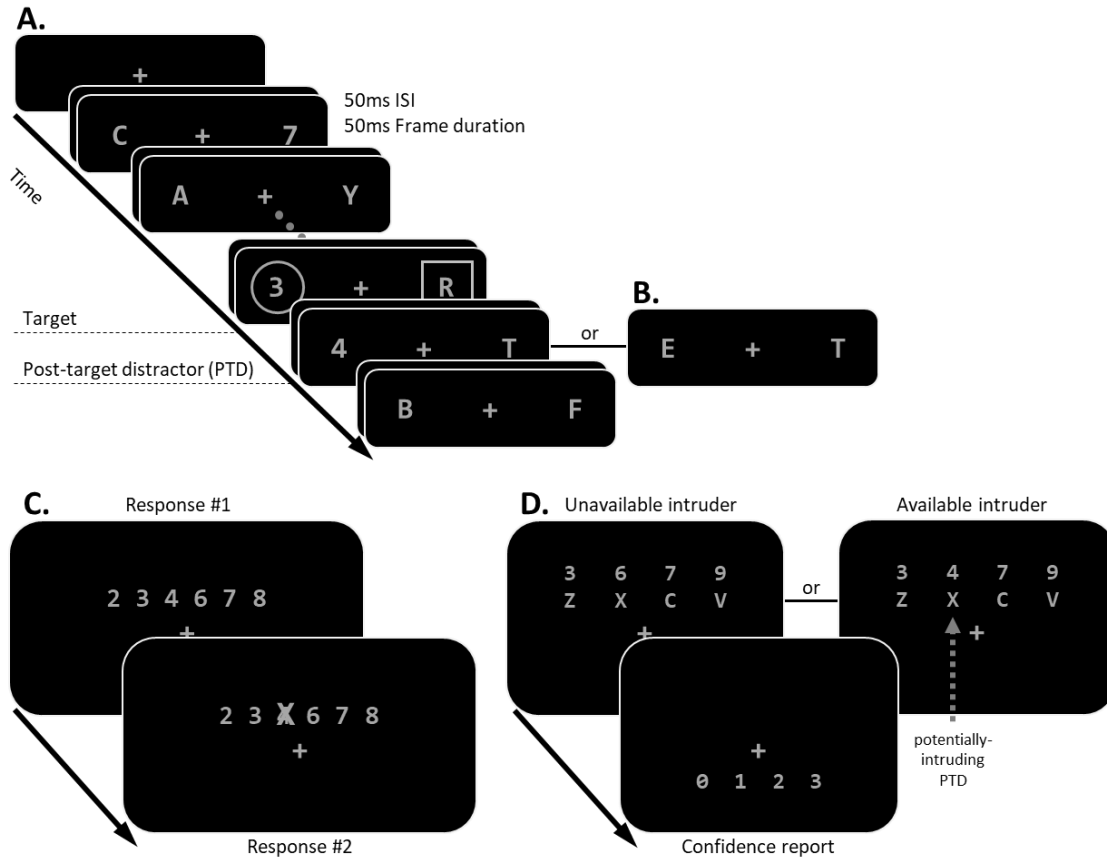


Figure 1. Illustration of the stimulus sequence in Experiment 1, 2 and 4. Participants had to report the target digit in one of two RSVP streams, defined by a pre-defined selection feature (e.g., circle). The target appeared at positions 5 to 8 within a stream, and was followed by three additional frames. At the same location as the target, the frame contained (A) a digit (i.e., category-matching) post-target distractor (PTD) on two thirds of trials and (B) a letter (i.e., category-nonmatching) PTD on one third of trials. C: Successive response screens presented in Experiment 1 at the end of each trial. In this example, the first response was “4”, which was crossed out in the second screen. D: Response screens in Experiment 2, containing only four possible response alternatives, which always included the target digit. The potentially intruding PTD was present in the half of all trials and absent in the other half. Each response screen was followed by a confidence report screen.

Results

First response

A preliminary analysis indicated that the shape of the selection feature (square vs. circle, varied across participants) did not affect accuracy rates, $F(1,14) = 1.33$, $p = .27$, $\eta^2_p = .09$, and data were

therefore combined across all participants. As expected, there was a strong distractor intrusion effect (Figure 2A). Accuracy on the first response was lower on trials where the target was followed by a potentially intruding PTD ($M = 40.8\%$) than when it was followed by a category-nonmatching PTD ($M = 69.8\%$), $F(1,15) = 57.84$, $p < .001$, $\eta^2_p = .79$. On 44.9% of all trials with a potentially intruding PTD, this distractor was reported on the first response (accounting for 75.8% of all errors on these trials).

Second response

Calculation of second responses rates included only trials that contained a potentially intruding PTD. When their first response choice was correct, participants reported the PTD as the second alternative on 44.3% of all trials. When they had reported the PTD as their first choice, they picked the target digit on 39.2% of all trials (Figure 2B). The small difference between these proportions was not statistically reliable, $t < 1$. Importantly, and in contrast with the prediction of the encoding competition account, both probabilities were significantly above chance (20%), $t(15) = 4.05$, $p = .001$, $d = 1.01$, and $t(15) = 4.59$, $p < .001$, $d = 1.15$, respectively, demonstrating that these second response choices were not random guesses. At the same time, and in contrast with the predictions of the post-encoding competition account, these two conditional probabilities were both reliably lower than the probability of correctly reporting the target on baseline trials with a category-nonmatching PTD (69.8%), $t(15) = 6.33$, $p < .001$, $d = 1.58$, and $t(15) = 4.17$, $p < .001$, $d = 1.04$, respectively¹. This result indicates that the target and the PTD competed for access to WM, reducing the probability that either item was reported relative to trials where no encoding

¹ An alternative way of testing the post-encoding competition account is to compare accuracy rates for first responses on baseline trials with the probability of target reports in either the first or the second response on trials with category-matching PTDs. To take account of fortunate guesses on trials where the target was not actually encoded, we first applied the lucky-guess correction procedure proposed by Olivers, Van der Stigchel, & Hulleman (2005). Following this correction, the estimated probabilities for actually encoding the target was 63.8% for baseline trials and 47.6% for trials with category-matching PTDs. This difference was significant, $t(15) = 3.32$, $p = .004$, $d = 0.81$, confirming that the presence of PTD digits affected target encoding.

competition was present. On the 14.3% of all trials where neither the target nor the potential intruder was picked as first choice, the probability of reporting the PTD as second choice was 34.7% and the probability of reporting the target was 27.3%. Both percentages were higher than chance, $t(15) = 3.82, p = .002, d = 0.96$, and $t(15) = 2.28, p = .04, d = 0.57$. The frequencies of each combination of first and second response choices are shown in Table 1, separately for trials with a category-matching PTD and baseline trials.

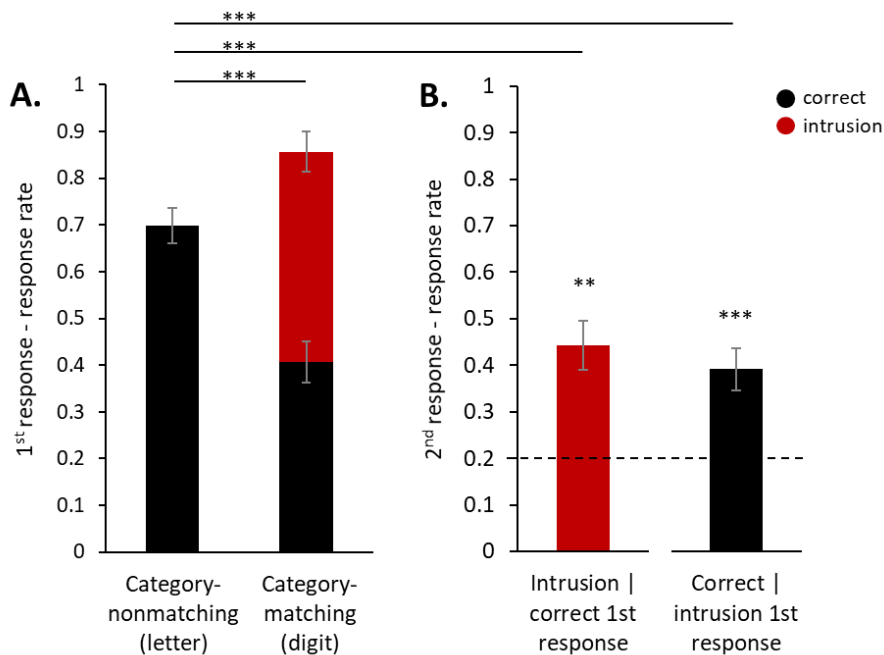


Figure 2. A: Frequency of correct reports and distractor intrusions errors for the first response choice on trials with category-nonmatching (letter) or category-matching (digit) post-target distractors. B: Frequency of second response choices on trials with category-matching post-target distractors (post-target distractor reports when the first response choice was correct; target reports when the first response was a distractor intrusion). The dotted line reflects chance levels. Error bars reflect one standard error.

Note. ** $p < .01$, *** $p < .001$

Table 1. *Frequency of first and second response choice combinations in Experiment 1, for trials with category-matching post-target distractors (left) and baseline trials (right).*

First response	Category-matching PTD				Category-nonmatching PTD (baseline)		
	Correct 2 nd	Intrusion 2 nd	Non-intrusion 2 nd	Total	Correct 2 nd	Non-intrusion 2 nd	Total
Correct 1 st	-	0.18	0.23	0.41	-	0.70	0.70
Intrusion 1 st	0.17	-	0.28	0.45	-	-	-
Non-intrusion 1 st	0.03	0.05	0.06	0.14	0.11	0.19	0.30
Total	0.20	0.23	0.57	1.00	0.11	0.89	1.00

Discussion

Experiment 1 produced two clear findings. First, as expected, the probability of correct target reports for the first response was much lower when the target letter was followed by a potentially intruding PTD than when it was target followed a category-nonmatching letter distractor (see Zivony & Eimer, in press, for analogous results). Second, we assessed the likelihood that one of two successive digits was encoded on trials where the other digit was also encoded, by measuring the conditional probabilities of correct target reports following a distractor intrusion response (correct | 1st intrusion), and distractor intrusions responses following correct reports (intrusion | 1st correct). Importantly, these conditional probabilities were not in line with the predictions of either the encoding or post-encoding competition accounts. They clearly exceeded the probabilities that would be expected if participants only ever successfully encoded one item per trial, and the second response was therefore chosen at random. This shows that in contrast to the encoding competition hypothesis, both items were encoded in WM on a subset of all trials. On the other hand, these conditional probabilities were far lower than the probability of correct target reports on baseline trials with category-nonmatching post-target distractors. An additional analysis based on the overall probability of target reports across both reports (footnote 1) produced the same result. If there had been no competition for encoding between targets and PTDs at all, as proposed by the post-encoding competition account, these probabilities should not have differed.

These results suggest a modified version of the encoding competition account, according to which the competition between the potentially intruding PTD and the target *reduces* the likelihood that both of these items will be encoded, but does not *necessarily* block either of them from entering WM. However, there are several alternative ways of explaining the pattern of results observed in Experiment 1 as a result of competitive interactions at post-encoding stages. The next three experiments were conducted to test each of these accounts. Experiments 2 and 3 investigated whether these results may reflect post-encoding mechanisms that prevent the target or the PTD from being selected for perceptual reports. Experiment 2 tested whether the results of Experiment 1 reflect a response strategy or limited access to perceptual awareness. Experiment 3 tested whether competitive interactions in WM impairs the precision of target and PTD representations in WM. Finally, Experiment 4 employed electrophysiological markers to test whether both of these items are initially encoded, but one is subsequently dropped from WM.

Experiment 2

In Experiment 1, participants reported both the target and the PTD much less frequently than they reported the target on baseline trials, suggesting that on many trials, only one of these items was encoded to WM. The alternative post-encoding competition account assumes both entered WM, and that subsequent competitive processes within WM often prevented one of them from being selected for perceptual reports. Experiment 2 investigated two possible competitive interactions that could have such an effect. It is possible that while both the target and the PTD were encoded, only one of them was strongly represented as appearing simultaneously with the target-defining selection feature on any given trial. As they were instructed to detect the digit that coincided with a particular shape, participants would have readily reported the item that they

experienced to have matched this target definition. However, they may have been reluctant to also report the identity of the other encoded item after specifying the perceived target, because they were confident that this item was not the target. Such a raised response threshold for reporting the other item held in WM could explain the results of Experiment 1.

To test this, we introduced a new manipulation in Experiment 2. On each trial, participants gave only one response out of a limited subset of four possible items which included the target (Figure 1D). The critical new manipulation was whether the response screen also included the potentially intruding PTD or not (intruder-available and intruder-unavailable trials). We reasoned that on intruder-unavailable trials, participants would be encouraged to correctly pick the target when it was encoded in WM, as it would be the most promising candidate, even if it was not strongly represented to coincide with the selection feature. Both our modified encoding competition account and the post-encoding competition account assume that the target and the PTD are encoded in WM, at least on some trials. For this reason, both accounts predict that correct target reports will be more frequent on intruder-unavailable relative to intruder-available trials, as the absence of the potentially intruding PTD from the set of response options should increase the probability of picking the target when it is encoded in WM. However, if there was no competition for WM encoding at all between the target and the PTD (as assumed by the response threshold hypothesis), accuracy on intruder-unavailable trials should be similar to baseline trials (target digit followed by a letter distractor), as in both types of trials, only one item that was available for report was also stored in WM. In contrast, if competition between the target and PTD reduces the likelihood that the target will gain access to WM, correct target reports should be considerably less frequent on intruder-unavailable trials relative to baseline trials.

A second possibility is that while both the target and the PTD are encoded in WM, competitive

interactions between them result in only one of them being consciously perceived on numerous trials. This ‘experiential blink’ account was supported by the results of an attentional blink study by Pincham et al. (2016) who found that visual awareness of one of two targets was reduced at lag-1, despite overall high accuracy for both targets. These results suggest that these two items did not compete for access to WM, but instead for access to conscious perception. To test this experiential blink account, participants in Experiment 2 were asked at the end of each trial to rate their confidence in their perceptual report on a four-point scale from “complete guess” to “full confidence”. Such confidence judgments are interpreted as an index of the ability to monitor the accuracy of visual representations (Mamassian, 2016), and usually strongly correlate with the amount of available perceptual evidence about a target (e.g., Desender et al., 2018; Guggenmos et al., 2016). However, confidence and accuracy have been shown to be dissociable in certain occasions. That is, previous studies have explored conditions where participants were either consistently over-confident in their inaccurate responses or under-confident in their correct responses (Baranski & Petrusic, 1994; Gigerenzer et al., 1991; Harvey, 1997).

The inclusion of confidence ratings in Experiment 2 allowed us to test a prediction that differentiates between the modified encoding competition account and the alternative experiential blink account. Both accounts suggest that on baseline trials with category-nonmatching PTDs, confidence judgements should be well calibrated, with target detection accuracy close to (though not necessarily at) chance level on trials where participants reported that they were guessing (i.e., when they declared “0” confidence in their response). However and importantly, the two accounts differ in their predictions regarding the relation between confidence and accuracy in intruder-unavailable trials. The encoding competition account assumes that confidence and accuracy are not dissociated in these trials, as participants either have clear perceptual evidence

(when the target is encoded in WM) or no perceptual evidence (when the target is not encoded) to guide their response selection. Because low confidence reports correctly reflect the absence of a target representation in WM on both baseline and intruder-unavailable trials, accuracy should therefore not differ between these trials. In contrast, the experiential blink account predicts a dissociation between accuracy and confidence, specifically for intruder-unavailable trials. If the target was often represented in WM yet blocked from awareness by the simultaneous presence of the PTD in WM, the target should frequently be selected for report when the intruder is unavailable, even when participants remain unaware of the target and thus have minimal confidence in their response selection. The experiential blink account therefore predicts not only above-chance accuracy on intruder-unavailable trials where confidence is low, but also higher accuracy on these trials relative to low-confidence baseline trials, where no distractor blocks access to awareness.

Method

For Experiments 2 and 3, we used a similar approach as Experiment 1 and calculated the sample size required to the weakest possible effect of interest when comparing accuracy in the new unavailable intruder condition to (i) the baseline condition and to (ii) the available intruder condition. Based on the data from Zivony and Eimer (in press, Experiment 1A) we made the following estimates. Accuracy in the baseline condition was again estimated to be 77%. Accuracy in the available intruder condition was estimated to be 36% (based on accuracy in the digit distractor condition in this previous experiment). The mid-point accuracy between these two conditions, giving rise to the smallest possible differences of interest, was therefore 56.5%. Error variance was again estimated based on the comparison of accuracy rates between digit distractor

and letter distractor conditions ($Sd = 12.5\%$). A power analysis with G*Power (Faul et al., 2013), using an alpha of .05 and power of .80, showed that the minimum sample size required to obtain reliable effects was 6 participants. Once again, due to the design differences between these experiments, we opted to use a larger sample of 16 participants.

Participants

Participants were 16 (12 women) volunteers ($Mage = 22.6$, $SD = 4.1$) who participated for £5. All reported normal or corrected-to-normal visual acuity.

Apparatus, stimuli and design

The apparatus, stimuli and design were identical to Experiment 1 expect for the following changes. To increase the number of possible responses, the set of possible targets was increased to include all digits from 2 to 9. At the end of each trial, two response screens were presented sequentially (Figure 1D). The first response screen showed four digits from which participants had to choose. These digits were presented 2.5° above fixation with an inter-item distance of 1.6° , sorted from left to right according to their numerical value (smallest digit on the left, largest digit on the right). The response screen also included the letters “Z”, “X”, “C”, and “V”, which appeared 1.5° above fixation, and were vertically aligned with the four digits. These letters specified the response keys assigned to each of the digits shown. One of these digits was always the target. On half of all trials with category-matching PTD (i.e, potential intruders), this item was also present on the response screen, while the other two were randomly drawn from the set of remaining digits (intruder-available trials). On the other half, the possible post-target intruder was not included in this response screen, which showed the target digit among three other randomly chosen digits

(intruder-unavailable trials). Participants chose one of the digits on the response screen by pressing the corresponding key with the index, middle, ring, or little finger of the left hand. The second response screen prompted participants to report their confidence regarding the target choice on this trial. This screen included the numerals “0”, “1”, “2”, and “3” that appeared 1.6° below fixation with 1° center-to-center inter-item distance. Confidence ratings ranged from 0 (“complete guess”) to 3 (“very high confidence”), and confidence judgments were made by pressing the corresponding number key with the thumb, index, middle, or ring finger of the right hand.

Results

Accuracy and intrusion rates

Once again, preliminary analysis indicated that the shape of the selection feature (square vs. circle) had no effect on accuracy rates, $F < 1$, and therefore data were collapsed across all participants. As can be seen from Figure 3A, accuracy was higher on trials with a category-nonmatching PTD ($M = 80.3\%$) relative to trials with a potentially intruding PTD that was available for report ($M = 50.9\%$), and this difference was significant, $t(1,15) = 8.93$, $p < .001$, $d = 2.23$. On intruder-available trials, 41.7% of the responses were distractor intrusions (representing 85% of all errors on these trials).

Notably, the percentage of correct target reports on intruder-unavailable trials was higher relative to intruder-available trials ($M = 68.9\%$ versus 50.9% ; Figure 3A), $t(1,15) = 11.53$, $p < .001$, $d = 2.88$. However, this difference could simply be due to random guessing on distractor intrusion trials where the PTD was perceived but was not among the four digits on the response screen. On these trials, the probability that the target is selected randomly is 25%. The frequency that the PTD was perceived on intruder-unavailable trials should be the same as on intruder-available trials (41.7%), as these two types of trials were randomly intermixed and

physically identical prior to the presentation of the response screen. Random guessing should therefore increase the rate of correct responses on intruder-unavailable trials by 10.4% ($25\% \times 41.7\%$), resulting in a guessing-adjusted expected accuracy rate of 61.3% ($50.9\% + 10.4\%$). Importantly, this rate was still significantly lower than the observed accuracy for intruder-unavailable trials, $t(1,15) = 6.65, p < .001, d = 1.66$, demonstrating that target digits were picked more frequently on these trials than would be expected on the basis of random guessing alone. However, and equally important, accuracy on intruder-unavailable trials was significantly lower relative to baseline trials with a category-nonmatching (letter) PTD (68.9% versus 80.3%), $t(1,15) = 5.27, p < .001, d = 1.32$. Thus, the presence of a potentially intruding PTD impaired participants' ability to access target identity even when this distractor was not available for report.

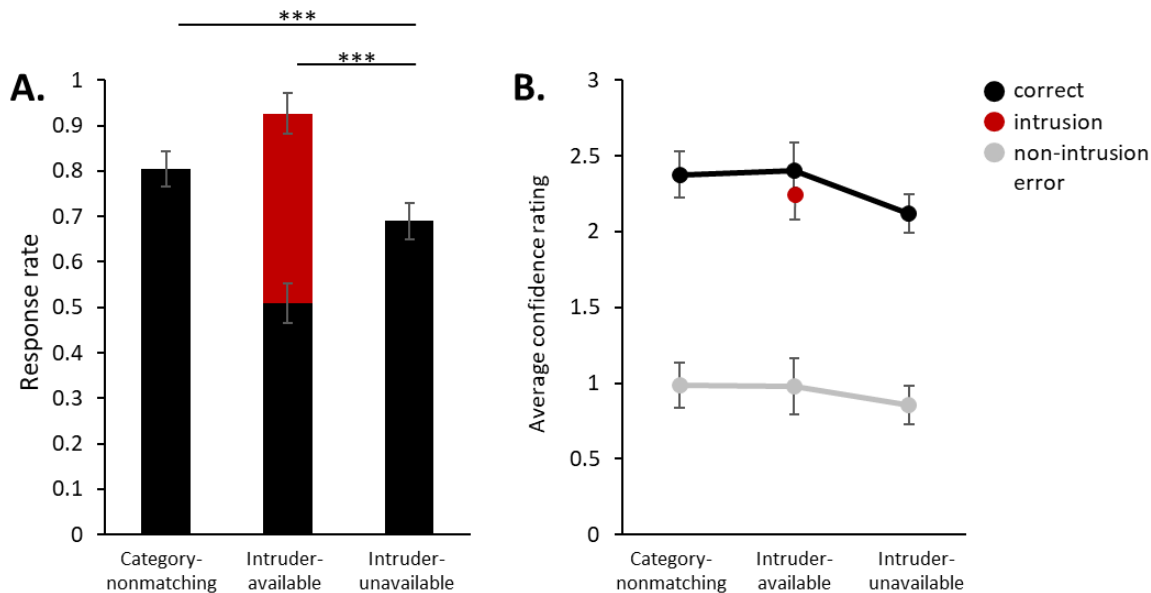


Figure 3. Mean response rates and confidence ratings in Experiment 2 on trials with category-nonmatching post-target distractors (i.e. letters), and the two types of trials with category-matching post-target distractors (intruder-available and intruder-unavailable). **A:** Frequency of correct responses and distractor intrusions. **B:** confidence ratings following correct responses, distractor intrusion, and non-intrusion errors. Error bars reflect one standard error. *Note.* *** $p < .001$.

Confidence ratings

Figure 3B shows average confidence ratings on trials with category-nonmatching PTDs and the two types of trials with potentially intruding PTDs (intruder-available and intruder-unavailable), separately for correct target reports (black line) and trials with non-intrusion errors where a nontarget digit other than the PTD was reported (grey line). The mean confidence rating following PTD intrusion errors is shown separately (in red). As would be expected, confidence was higher following correct responses ($M = 2.30$) relative to non-intrusion errors ($M = 0.91$), $t(15) = 12.02$, $p < .001$, $d = 3.01$. Notably, confidence ratings following distractor intrusions on intruder-available trials ($M = 2.19$) were also much higher than on trials with non-intrusion errors, $t(15) = 9.58$, $p < .001$, $d = 2.40$. However, confidence in distractor intrusion reports was slightly but significantly lower than confidence in correct target reports on intruder-available trials, $t(15) = 2.89$, $p = .01$, $d = 0.72$.

For confidence ratings following correct responses, a main effect of trial type was found, $F(1,15) = 19.81$, $p < .001$. Confidence did not differ between trials with category-nonmatching PTD and intruder-available trials ($t < 1$), but was significantly reduced on intruder-unavailable trials [$M = 2.37$ and $M = 2.40$ vs. $M = 2.12$; $t(15) = 5.90$, $p < .001$ and $t(15) = 4.32$, $p = .002$, following post-hoc Bonferroni corrections]. The distribution of confidence ratings following correct responses for all three types of trials is shown in Table 2. Relative to baseline trials with category-nonmatching PTD, confidence ratings on intruder-unavailable trials were mainly characterized by an increase in guessing (“0” ratings) and a decrease in full confidence (“3” ratings; Table 2, rightmost column),

Table 2. Frequency of confidence responses on correct trials as a function of post-target distractor condition (category-nonmatching, intruder-available, intruder-unavailable trials). The rightmost columns show the results of analyses comparing the rates of specific confidence responses between trials with category-nonmatching post-target distractors and intruder-unavailable trials.

Confidence rating	Category-nonmatching	Available intruder	Unavailable intruder	Unavailable intruder vs. category-nonmatching
0	5%	2%	11%	$\bar{d} = 6\%$, $t(15) = 3.27$, $p = .005$, $d = 0.82$
1	12%	10%	13%	$\bar{d} = -1\%$, $t(15) = 0.83$, $p = .42$, $d = 0.21$
2	25%	33%	29%	$\bar{d} = 4\%$, $t(15) = 1.44$, $p = .17$, $d = 0.36$
3	58%	55%	47%	$\bar{d} = -11\%$, $t(15) = 5.80$, $p < .001$, $d = 1.45$

It is plausible to interpret “0” confidence reports following correct responses on intruder-unavailable and baseline trials as evidence that targets were not encoded into working memory on these trials, and the correct response was picked as a result of a lucky guess. In this case, the increase of “0” confidence ratings on intruder-unavailable trials would show that targets were less likely to gain access to working memory on these trials relative to baseline trials with a category-nonmatching PTD. Alternatively, as suggested above, it is possible that on some intruder-unavailable trials the target was encoded into working memory and this representation was then used to choose the correct response, but participants remained metacognitively unaware of this fact, as reflected by “0” confidence ratings. If this was correct, accuracy rates when confidence was minimal should be significantly higher on intruder-unavailable trials relative to baseline trials. However, this was not the case. The rate of correct responses that were followed by a “0” confidence judgment was nearly identical on intruder-unavailable trials and on baseline trials with category-nonmatching post-target distractors (30.5% versus 33.4%; $t < 1$) demonstrating that participants’ expressed confidence accurately reflected the perceptual evidence they had about the target’s identity. Since the absence of a significant effect does not constitute as evidence in favour of the null hypothesis, we also calculated the Bayes Factor associated with the difference between these trials and interpreted the strength of the evidence based on the classification suggested by

Jeffreys (1961). We conducted a one-sided dependent-sample Bayesian t-test using JASP (0.9.2) with the default prior of Cauchy scale of 0.707, and found support for the null hypothesis, i.e., that there is no difference in average accuracy between these conditions, $BF_{01} = 5.76$,

Discussion

The results of Experiment 2 provided new evidence for our modified encoding competition hypothesis, and against two alternative post-encoding competition accounts. As in Experiment 1, there was again a large number of PTD intrusions when participants could report the PTD (intruder-available trials). Notably, on trials where the potentially intruding PTD was unavailable for report, accuracy was higher than on intruder-available trials. This observation provides further evidence that the target was encoded in WM and was accessible to report on at least some trials where participants would have reported the PTD if it had been available. However, and importantly, accuracy on intruder-unavailable trials was still much lower than on baseline trials. This is not in line with a response threshold explanation of the results of Experiment 1, which assumes that participants are reluctant to report a second item that is represented in WM after having reported their first choice. In this case, the frequency of target reports on distractor-unavailable trials and baseline trials should have been similar. The fact that this was not the case suggests that the target was not encoded in WM on a proportion of the trials where a PTD digit was present.

In addition, Experiment 2 found no support for the experiential blink explanation of the PTD effect, which assumes that the target and the PTD are both represented in WM, but one of them is often blocked from access to conscious awareness (Pincham et al., 2016). Relative to baseline trials, high-confidence responses decreased and guessing responses increased on

intruder-unavailable trials. However, and importantly, when participants reported that they were guessing, accuracy was equally low in both types of trials. This suggests that the decrease in confidence on intruder-unavailable trials was not due to target representations in WM failing to become available for conscious reports. Instead, participants appeared to have been able to accurately monitor the perceptual evidence (or its absence) that guided their response selection.² Overall, the findings of Experiment 2 provides more evidence that competition between the target and the potentially intruding PTD reduces the likelihood that the target will be encoded in WM.

Finally, Experiment 2 showed that distractor intrusion errors were associated with a high degree of confidence (see also Recht et al., 2019, for similar results). This suggests that participants based their response choice on these trials on strong perceptual evidence that the PTD digit appeared at the same time as the selection feature. Nevertheless, confidence in these intrusion responses was slightly but reliably lower than confidence in correct target reports. This could be related to attentional engagement being slower on trials with intrusion responses relative to correct responses, as revealed by our previous N2pc results (Zivony & Eimer, in press).

Experiment 3

The findings of Experiments 1 and 2 suggests that the competition between the target and PTD can prevent the target from being encoded in WM. However, these results could also be explained by another alternative post-encoding competition account, which assumes that competitive interactions between these two items reduces the precision of their representations in WM (see

² It should be noted that Pincham et al. (2016) asked participants to rate their subjective awareness of a target rather than their confidence. These two types of ratings have been argued to be dissociable (e.g., Jachs et al., 2015; Overgaard et al., 2010), as one may have a visual experience of the target but have little to no confidence that this experience contributed to accurate reports. However, accurate calibration between low confidence and low accuracy can only when participants have little-to-no visual awareness of the target. Therefore, the fact that low-confidence reports in Experiment 2 were equally associated with low accuracy on baseline and intruder-unavailable trials strongly suggest that these reports does reflect a lack of visual awareness that is the result of the target not having been encoded in WM.

Fougnie et al., 2012, for similar ideas). If the target's representation in WM was strongly degraded on some trials, it may no longer be able to support the selection of the correct response.

To test this possibility, we added a new manipulation to the design of Experiment 2. On half of all trials in Experiment 3, the forced choice response screens included a non-target that was physically similar to the target (henceforth a '*lure*'). On the other half, no such lure was present. To be able to include lure stimuli, target stimuli were now chosen from a set of eight letters that could be paired to similar non-target letter (Miller & Bauer, 1981; Townsend, 1971; see Figure 4A and 5C). As in Experiment 2, there were baseline trials without a PTD letter (i.e., they included a category-nonmatching digit PTD), as well as trials with intruder-available and intruder-unavailable response screens. Lures were equally likely to be included in the response screens in all three types of trials. If competitive interactions between the target and PTD in WM reduce the precision of the representation of the target, participants should be more susceptible to choosing the lure over the target on trials with a potentially intruding distractor relative to baseline trials. Furthermore, this tendency might be even more pronounced on trials where the intruder is perceived as target but is unavailable for report, and response selection has to be based on a particularly degraded WM representation of the target. Alternatively, if the competition between the target and PTD operates at the encoding stage by preventing the target from entering WM on some trials, any effects of the lure on accuracy should reflect general limitations in the precision of target representations in WM that is unaffected by the presence of a PTD. Therefore, according to the encoding competition account, these effects should not differ between baseline trials, distractor-available trials, and distractor-unavailable trials.

Method

Participants

Participants were 16 (10 women) volunteers ($M_{age} = 23.5$, $SD = 2.0$) who participated for £7. All reported normal or corrected-to-normal visual acuity.

Apparatus, stimuli and design

The apparatus, stimuli and design was identical to Experiment 2 except for the following changes. First, the roles of digits and letters were reversed, as participants now had to report the identity of a target letter (instead of a digit). Accordingly, the post-target distractor was a letter (i.e., category-matching) on two thirds of all the trials and a digit (i.e., category-nonmatching) on one third of the trials (Figure 4A and 5B). The critical manipulation in Experiment 3 was whether a nontarget letter that was similar to the target letter (a “lure”) was included in the response screen or not. For this reason, a set of eight possible target letters was employed (“C”, “G”, “I”, “M”, “N”, “T”, “U”, and “V”), which included four pairs of letters that were physically similar to each other, but dissimilar from the other letters (Figure 4C). When a post-target letter distractor was present in the RSVP streams, this letter was always dissimilar to the preceding target. The response screen presented at the end of each trial contained four out of the potential target set of eight letters (Figure 4D). One of these letters was the target. On half of all trials, one of the other letters was similar to the target and the two others were dissimilar (lure-present trials). On the other half, all three nontarget letters were dissimilar to the target (lure-absent trials). On half of all trials where a category-matching post-target letter distractor was presented, this letter was present in the response screen (intruder-available trials). On the other half, this letter was absent (intruder-unavailable trials). The letter that was similar to the post-target distractor letter never appeared in any response screen. The four letters were presented in their alphabetical order from

left to right. Presented 1° below each letter were four digits, from 0 to 3, indicating the four key participants used to respond. In contrast to Experiment 2, no confidence judgments were required, and each trial ended after one letter from the response screen was chosen. Participants completed 10 practice trials, followed by 600 experimental trials presented in blocks of 50 trials.

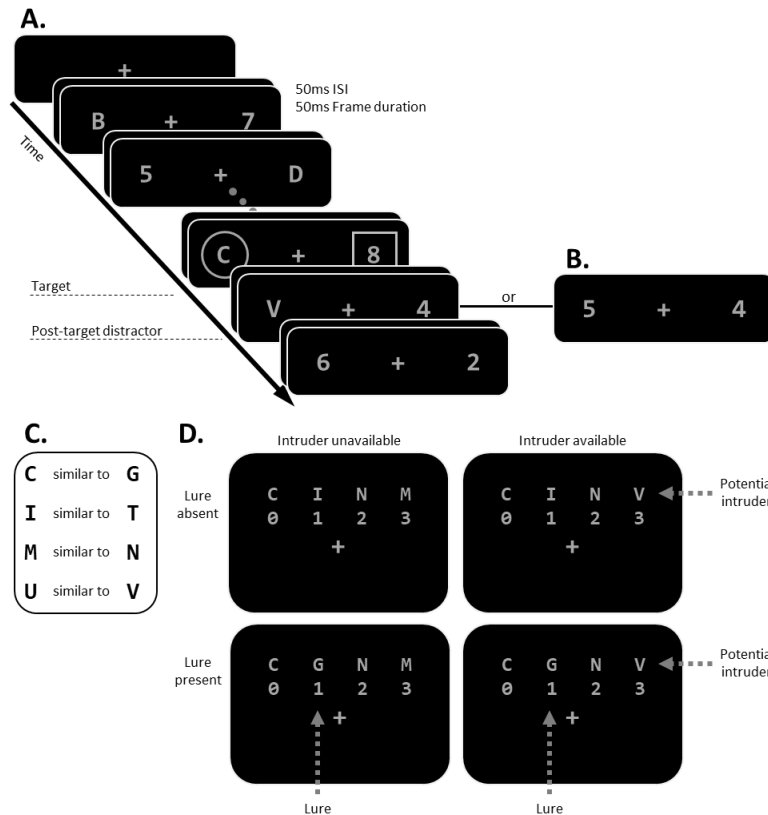


Figure 4. Illustration of the stimulus sequence in Experiment 3. Participants had to report the target letter. The post-target frame contained (A) a letter distractor at the same location as the target on two thirds of trials and (B) a category-nonmatching (digit) distractor on one third. C: The set of possible targets contained four pairs of similar letters. D: The four types of response screens (left/right: intruder-available versus intruder-unavailable; top/bottom: lure-absent versus lure-present).

Results

Once again, the shape of the selection feature had no effect on accuracy ($F < 1$), and data were combined across all participants. In an initial analysis, accuracy rates were collapsed across trials where a lure was present or absent in the response screen. As in Experiments 1 and 2, accuracy was

higher on trials where the post-target letter distractor did not match the target category (i.e., a digit in Experiment 3) relative to trials where it matched this category (i.e. a letter) and this distractor was available for report ($M = 71.9\%$ versus 43.8% ; $t(15) = 12.58$, $p < .001$, $d = 3.15$). On intruder-available trials, 44.6% of all responses were distractor intrusions (representing 79.3% of all errors on these trials). Analogous to Experiment 2, accuracy on intruder-unavailable trials ($M = 56.5\%$) was significantly higher than on intruder-available trials, $t(15) = 8.04$, $p < .001$, $d = 1.01$, but lower than on trials with a category-nonmatching post-target distractor, $t(15) = 6.26$, $p < .001$, $d = 1.57$.

To assess the impact of the presence versus absence of a lure in the response screen (labelled P versus A in Figure 5) for the three different types of trials, we conducted an ANOVA with the factors post-target distractor trial type (category-nonmatching, intruder-available, and intruder-unavailable) and response screen (lure-present, lure-absent). There was a small but reliable general reduction in accuracy on trials where a lure was included as one of the response options relative to trials where it was not included in the response screen, $M = 55.1\%$ vs. $M = 59.6\%$, $F(1,15) = 10.64$, $p = .005$, $\eta^2_p = .415$. However, and importantly, the interaction between the two factors was not significant, $F < 1$, suggesting that the accuracy costs produced by making a target-similar lure item available for report were not modulated by the presence versus absence of a post-target category-matching distractor in the preceding RSVP streams. We also calculated the Bayes Factor associated with the interaction by entering the two factors into a repeated measures Bayesian ANOVA using JASP (0.9.2) with the default prior of $r_A = 0.5$. The full model (including the interaction and both main effects, $BF_{10} = 1.31e+22$) was then divided by the model associated with the two main effects ($BF_{10} = 7.95e+22$), in order to isolate the contribution of the interaction to the model (Wagenmakers et al., 2018). This analysis provided strong support for the null

hypothesis that the accuracy costs produced by including a lure in the response screen were identical across the three types of trials, $BF_{01} = 6.08$. The frequency of intrusion errors on intruder-available trials was similarly not affected by the presence versus absence of a lure in the response screen, $t < 1$.

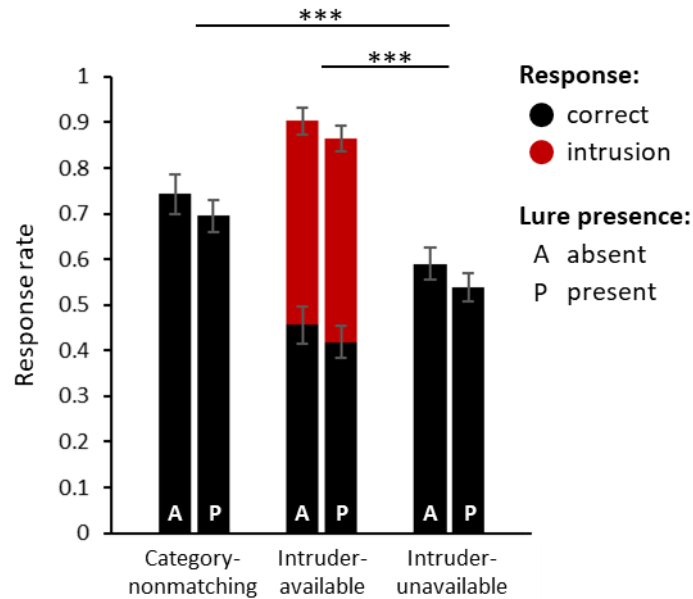


Figure 5. Frequency of correct responses and distractor intrusions as a function of trial type (category-nonmatching, intruder-available, intruder-unavailable) and response screen type (lure absent versus present: A versus P). Error bars reflect one standard error. Note. *** $p < .001$

Discussion

Experiment 3 yielded two clear cut results. First, overall accuracy results confirmed those of Experiment 2. Accuracy was highest on baseline trials, lower on intruder-unavailable trials, and lower still on intruder-available trials. Second, the inclusion of a lure that was physically similar to the target as one of the response options reduced overall accuracy. However, and critically, these lure-induced costs were virtually identical on all three types of trials. If competition between the target and the PTD in WM had reduced the precision of the target representation, the presence of a

lure should have reduced report accuracy much more strongly for trials where both items were presented relative to baseline trials. The fact that no such difference was found in Experiment 3 provides clear evidence against this version of a post-encoding competition account, but is entirely consistent with the view that this competition blocks the target from entering WM on some trials. The general lure-induced costs observed in Experiment 3 indicates that, on a small proportion of trials, the precision of the target's representation in WM was insufficient to discriminate between the target letter and the physically similar letter, regardless of whether the target was followed by a potentially intruding PTD or not.

Experiment 4

In this final experiment, we sought additional and more direct evidence that competitive interactions between the target and PTD can act prior to encoding at perceptual stages, resulting in only one of these two items gaining access to WM on a substantial number of trials. So far, this conclusion was based entirely on behavioural results from perceptual reports and confidence ratings. In Experiment 4, we employed the two-response procedure introduced in Experiment 1, but now also measured event-related potentials (ERPs) as direct on-line markers of WM storage. We focused on the contralateral delayed activity (CDA), which is an established electrophysiological index of WM storage (see Luria et al., 2016, for review). The CDA is elicited during the delay period of lateralised WM tasks as an enhanced negativity at posterior electrodes contralateral to the side of to-be-memorized visual items. CDA amplitudes increase with the number of items that are stored in WM and are also sensitive to individual differences in WM capacity (e.g., Vogel & Machizawa, 2004; Ikkai et al., 2010; Luck & Vogel, 2013; McCollough et al., 2007), demonstrating that that they reflect neural mechanisms involved in the on-line

maintenance of visual information. In Experiment 4, we measured CDA components elicited after the presentation of a target item in the left or right stream. In contrast to the previous experiments, the target was followed by a category-matching (digit) PTD on all trials. Based on the results of Experiment 1, where participants reported both the target and the PTD on some trials, but only one of these items on other trials, we compared CDA amplitudes measured on these two types of trials. If both items are encoded in WM on trials where both are reported (two-items report trials), but only one item is stored on trials where either the target or the PTD (but not both) are reported (single-item report trials), this should be reflected by differences in CDA components, with larger CDA amplitudes in the former trials. In contrast, if single-item reports reflect occasions where the two items were encoded in WM but, for whatever reason, only one item was reported, there should be no difference in the CDA amplitude between single-item and two-item trials.

Measuring the time course of CDA components on these two types of trials also made it possible to test another post-encoding competition account that would be consistent with the results reported so far. It is possible that participants do in fact encode both the target and PTD in WM on the majority of trials, but are able to retain only one object long enough to report it. According to this “catch-and-release” account, one of the two encoded items is rapidly dropped from WM on trials with single-item reports, either due to competitive interactions with the other item, or because participants know that there is only a single target on each trial. If this hypothesis is correct, it should be reflected by systematic differences in the time course of CDA components on these trials as compared to trials where both the target and the PTD are reported. CDA components should initially be of equal size on both types of trials (reflecting the initial encoding and maintenance of both items in WM). Subsequently, the CDA amplitude for trials with single-item reports should rapidly decline relative to trials with two-items reports, reflecting the

release of one of these items from WM. Using the CDA to test the catch-and-release account assumes that this component is sensitive to such a process, and previous research (Balaban et al., 2018; Berggren & Eimer, 2016) has shown that CDA amplitude changes across time can reflect such fast changes in WM load³. In contrast, if only one of these two items is stored in WM on single-item report trials, as proposed by the encoding competition account, CDA amplitude differences between two-items and single-item report trials should be present from the moment when this component emerges (reflecting the presence of two items versus one in WM), and this difference should remain constant during the entire maintenance period.

In addition to the CDA, we also measured N2pc components to targets in Experiment 4, separately for trials where participants reported either the target or the PTD as their first choice. By comparing N2pc onset latencies between these trials, we aimed to confirm the N2pc onset delays for trials with distractor intrusion errors observed in our previous study (Zivony & Eimer, in press), indicating that these errors are associated with slower attentional engagement.

Method

Sample size selection

As this is the first distractor intrusion experiment to examine CDA amplitude differences between single-item response and two-item responses, we could not calculate the sample size required to find this effect. Since another aim of Experiment 4 was to confirm the N2pc latency delay for trials with distractor intrusions relative to trials with correct responses observed in our previous study (Zivony & Eimer, in press), we therefore based our sample size calculation on this previous effect. To do so, we combined the results of Experiment 1A (N=12) and 1B (N=11) and

³ For example, Balaban et al. (2018) found that when a tracked moving object split into two independent parts, CDA amplitude dropped sharply about 200 ms after this separation occurred, indicating that this object was temporarily dropped from WM (“memory resetting”).

calculated the N2pc onset latency effect, $F_{adjusted}(1,22) = 20.17$, $p < .001$, and the associated effect size, $\eta_p^2 = .48$. Because the onset latency analysis was based on jackknifed N2pc waveforms, it is questionable whether the effect size (as reflected by η_p^2) is meaningful in any context other than determining sample size for a similar analysis. Nevertheless, based on these data, we conducted a power analysis with G*Power (Faul et al., 2013), using an alpha of .05, and power of .80. This analysis revealed that the minimum sample size required to obtain a reliable effect was 12 participants. For sake of comparison with Experiments 1-3, and in order to maximize the chances to also observe reliable CDA amplitude differences, we used a sample of 16 participants in Experiment 4.

Participants

Participants were 16 (11 women) volunteers ($M_{age} = 26.63$, $SD = 9.31$) who participated for £25. All reported normal or corrected-to-normal visual acuity. One participant was excluded from all analysis because of excessive eye movement and eye blinks that resulted in rejection of more than 75% of their EEG data.

Apparatus, stimuli and design

The apparatus, stimuli and design for identical to Experiment 1 expect for the following changes. We increased the number of possible response alternatives, in order to reduce the likelihood of correct response and intrusion responses occurring by chance. The set of possible targets was therefore increased to include all digits from 2 to 9. There were no baseline trials, as the target was followed by a category-matching PTD (i.e., another digit) on all trials. As in Experiment 1, participants were explicitly told that (i) there was only a single target, (ii) their task was to identify this target, and (iii) the purpose of the second response was to maximize the likelihood of

accurate target reports. To enable the measurement of CDA components during the retention phase and prior to response selection and execution, the response screen was preceded by a fixation display that was presented for 500 ms. The experiment included 20 practice trials followed by 600 experimental trials, divided into 50-trial blocks.

EEG Recording and Data Analysis

EEG was DC-recorded from 27 scalp electrodes, mounted on an elastic cap at sites Fpz, F7, F8, F3, F4, Fz, FC5, FC6, T7, T8, C3, C4, Cz, CP5, CP6, P9, P10, P7, P8, P3, P4, Pz, PO7, PO8, PO9, PO10, and Oz. A 500-Hz sampling rate with a 40 Hz low-pass filter was applied. Channels were referenced online to a left-earlobe electrode, and re-referenced offline to an average of both earlobes. No other filters were applied after EEG acquisition. Trials with eye blinks (exceeding $\pm 60 \mu\text{V}$ at Fpz), horizontal eye movements (exceeding $\pm 30 \mu\text{V}$ in the HEOG channels), and muscle movement artefacts (exceeding $\pm 80 \mu\text{V}$ at all other channels) were removed as artefacts. N2pc and CDA components were averaged separately, based on epochs starting 100 ms prior to the onset of the target frame, and ending 500 ms or 800 ms after frame onset, for the N2pc and CDA analyses respectively. The average loss of epochs due to artefacts prior to averaging was 14.2% ($SD = 10.4\%$) for the N2pc analysis and 29.1% ($SD = 20.3\%$) for the CDA analysis. There was no difference in the number of rejected epochs between the different experimental conditions, $F < 1$. All ERPs were averaged relative to a 100 ms pre-stimulus baseline. For both analyses, averaged ERP waveforms were computed for trials with a target in the left or right RSVP stream, in order to compare ERPs at electrodes PO7/PO8 contralateral and ipsilateral to the location of the target.

CDA. The CDA analysis focused on comparing trials where participants reported both the

target and the PTD (two-items report trials; irrespective of the order in which they were reported) to trials where participants reported either the target or the PTD (but not both), either on the first or second response (single-item report trials). The average number of epochs retained for analysis was $M = 256$ ($SD = 152$) for two-item report trials and $M = 141$ ($SD = 71$) for single-item report trials. The analysis window for CDA mean amplitudes was 400-800 ms after target frame onset. Although previous studies have used longer delay periods and longer CDA windows, the predictions tested in Experiment 4 were related to the initial period of the CDA. Moreover, the relatively shorter time window reduced data loss due to blinking. CDA amplitude was defined as the mean amplitude of the contralateral-ipsilateral difference waveform in the 400–800 ms time window after the onset of the target frame. CDAs were measured from 400 ms post-stimulus onwards in order to prevent any overlap of the CDA time window with the preceding N2pc component. In addition, CDA amplitudes were also quantified separately within a 400-500 ms interval (as the “catch-and-release” hypothesis predicts no CDA differences between single-item reports and two-item report trials in this early window), and a subsequent 500-800 ms interval. Because CDAs are reflected by negative values (i.e., contralateral negativities) in these difference waves, one-tailed t-tests against zero were used to assess the presence of CDA components within a specific time window on either single-item or two-items report trials.

N2pc. To confirm the observations from our previous study (Zivony & Eimer, in press), we compared the N2pc onset latencies and amplitudes between trials where the first response was correct and trials where participants chose the PTD for their first response. Trials where neither of these items was reported first were excluded. For this analysis, the identity of the second reported item was not taken into account, in order to maintain a sufficiently high signal-to-noise ratio for N2pc components. The average number of epochs retained for analysis was $M = 211$ ($SD = 122$)

for correct first-responses trials and $M = 232$ ($SD = 57$) for intrusion first-response trials. As in our previous study (Zivony & Eimer, in press), N2pc onset latencies were calculated on the basis of contralateral-ipsilateral difference waveforms, following an application of a 10Hz low pass filter (see also Brisson et al., 2007). We employed the jackknife procedure described by Miller et al. (1998), with the N2pc onset criterion defined as the point where the difference waveform reached 50% of the average N2pc peak amplitude (averaged across trials with correct responses and distractor intrusion trials, and measured within a 150-300 ms post-target interval). A relative onset criterion was used to avoid any distortions due to N2pc amplitude differences (see Zivony & Eimer, in press; Grubert & Eimer, 2015; Grubert et al., 2011, for similar procedures). In statistical analyses of N2pc onset latency differences, F scores were corrected according to the formula provided by Ulrich and Miller (2001). N2pc mean amplitudes were also compared between trials with correct responses versus distractor intrusions, based on mean amplitudes of ipsilateral-contralateral difference waveforms in the 200–300 ms time window after the onset of the target frame (e.g., Berggren & Eimer, 2019; Kiss et al., 2008; Zivony & Eimer, in press).

Results

Behavioural results

First response. A preliminary analysis indicated that the shape of the selection feature (square vs. circle, varied across participants) did not affect accuracy rates, $t(14) = 1.12$, $p = .28$, $d = 0.29$, and data were therefore combined across all participants. As shown in Figure 6A, the target was reported as the first response on 40.1% and the PTD was reported on 46.9% of the trials (accounting for 78.3% of error trials).

Second response. When their first response choice was correct, participants chose the PTD as

the second alternative on 70.7% of all trials. When they had reported the post-target distractor as their first choice, participants picked the target digit on 61.1% of all trials (Figure 6B). Both probabilities were significantly above chance (i.e., 14.3%), $t(14) = 10.55, p < .001, d = 2.72$, and $t(14) = 8.71, p < .001, d = 2.25$, respectively, demonstrating that these second response choices were not random guesses. In contrast to Experiment 1, intrusions errors following a correct response were significantly more frequent than correct responses following an intrusion error, $t(14) = 3.50, p = .003, d = 0.90$. On the 13.0% of all trials where neither the target nor the PTD was picked as first choice, the probability of reporting the PTD as second choice was 38.8% and the probability of reporting the target was 24.7%. Both percentages were higher than chance, $t(14) = 8.28, p < .001, d = 2.14$ and $t(14) = 3.97, p < .001, d = 1.03$. The frequencies for each combination of first and second response choices are shown in Table 3.

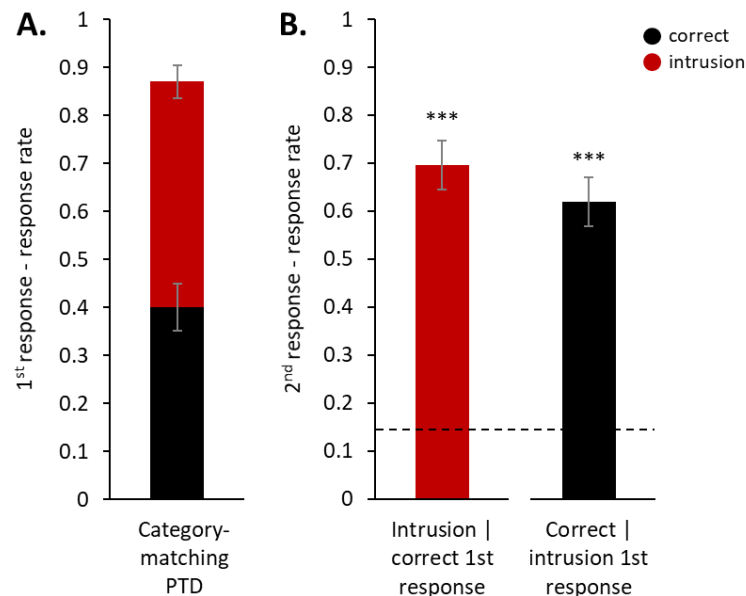


Figure 6. A: Frequency of correct reports and distractor intrusions errors for the first response choice in Experiment 4. B: Frequency of correct responses and distractor intrusions for the second response choices on trials where the first response was either an intrusion or correct, respectively. The dotted line reflects chance levels. Error bars reflect one standard error.

Note. *** $p < .001$

Table 3. Frequency of first and second response choice combinations in Experiment 4.

First response	Second response			Total
	Correct 2 nd	Intrusion 2 nd	Non-intrusion 2 nd	
Correct 1 st	-	0.30	0.10	0.40
Intrusion 1 st	0.27	-	0.20	0.47
Non-intrusion 1 st	0.03	0.04	0.06	0.13
Total	0.30	0.34	0.36	1.00

Electrophysiology

CDA components: Single-item versus two-items reports. Figure 7A shows the ERP waveforms triggered by the target frame at electrodes PO7 and PO8 contralateral and ipsilateral to the target in the 800 ms interval after target frame onset, separately for two-items and single-item response trials. The corresponding difference waves obtained by subtracting ipsilateral from contralateral ERPs are shown in Figure 7B. N2pc components were followed by clear CDA components on both types of trials. Notably, CDAs were smaller and appeared to emerge later on single-item as compared to two-items report trials. CDA mean amplitudes measured in the 400–800 ms time window were significantly different from zero on two-items report trials and also on single-item report trials, $p = .005$ and $p = .04$, respectively. Critically, mean CDA amplitudes were significantly larger on two-items report trials relative to single-item report trials ($M = -1.04 \mu\text{V}$ vs. $M = -0.53 \mu\text{V}$), $F(1,14) = 10.28$, $p = .006$, $\eta_p^2 = .42$.

As mentioned above, the key question that separates the encoding competition and catch-and-release accounts is whether CDA amplitude differences between single-item and two-items report trials are present from the start of the CDA, or only emerge at a later point in time. To examine this, we divided the overall CDA measurement interval into an early and later time window (400-500 ms and 500-800 ms after frame onset, respectively), and entered mean CDA amplitudes in a two-way ANOVA with trial type (single-item vs. two-items report) and time window (early, late) as independent variables. This analysis revealed a significant interaction between the two factors,

$F(1,14) = 6.05, p = .03, \eta_p^2 = .30$, reflecting the fact that CDA amplitude differences between single-item and two-items report trials were larger during the early time window (see Figure 8B)⁴. However, these differences were reliably present during both time windows (early: $F(1,14) = 14.77, p = .002, \eta_p^2 = .51$; late: $F(1,14) = 7.82, p = .014, \eta_p^2 = .36$). During the early window, the CDA was absent on single-item report trials, $t < 1$, but already present for two-items report trials, $t(14) = 1.94, p = .04, d = 0.50$. In the late time window, CDA amplitudes were reliably different from zero for both single-item and two-items report trials, $t(14) = 2.75, p = .008, d = 0.71$ and $t(14) = 3.92, p = .001, d = 1.01$, respectively.

⁴ It is possible that these CDA amplitude differences between two-item and single-item report trials were modulated by the identity of the item (target or PTD) that was reported first. To examine this possibility, we ran additional analyses where the identity of the first reported item was included as a factor. In these analyses, two participants who had less than 10 trials in one the four conditions remaining after artefact rejection were not included. No evidence for a systematic impact of which item was reported first was found. In the early time window, the CDA was larger on two-item report trials, $F(1,12) = 14.78, p = .002, \eta_p^2 = .55$, and on trials where the target was reported first, $F(1,12) = 9.29, p = .01, \eta_p^2 = .44$. However, these effects did not interact, $F < 1$. In the late CDA time window, the number of items affected the CDA, $F(1,12) = 5.94, p = .031, \eta_p^2 = .33$, whereas the identity of the first reported item did not, $F < 1$. Again, there was no interaction between these two factors, $F < 1$.

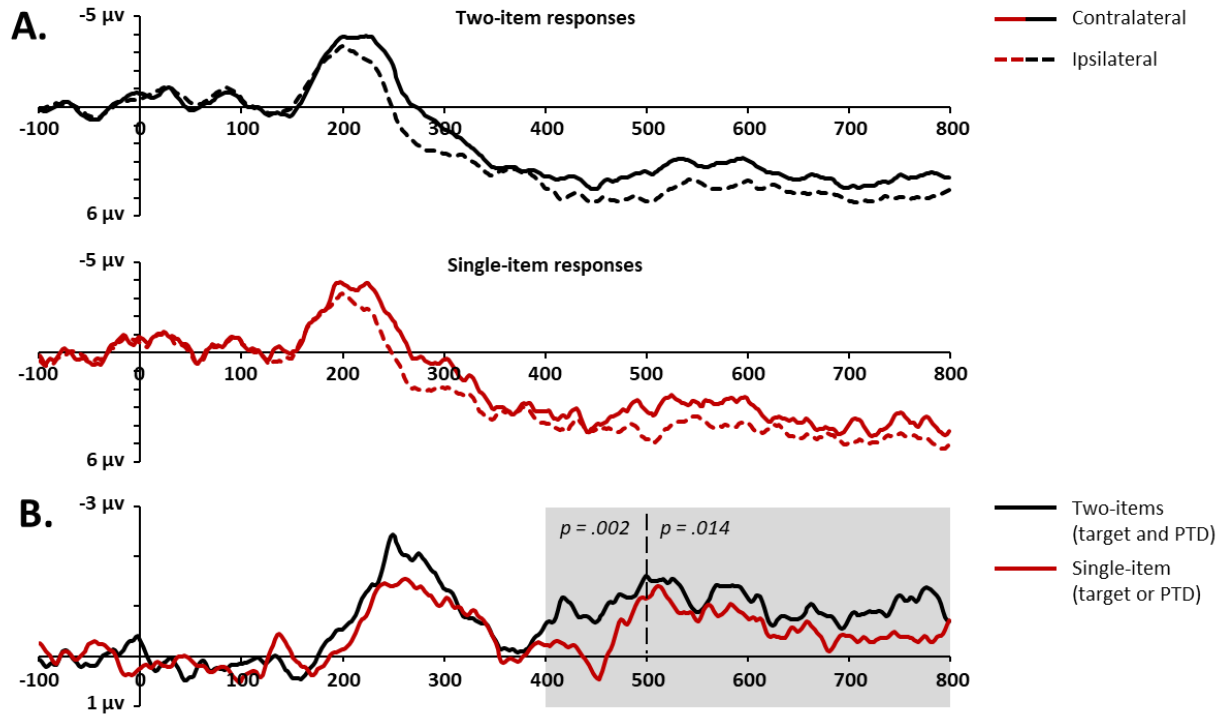


Figure 7. Grand-average event-related potentials (ERPs) waveforms on electrodes PO7/PO8 elicited in Experiment 4 by target frames, shown separately for trials where participants reported both the target and the PTD (two-items, black lines) and trials where participants reported either the target or the PTD, but not both (single-item, grey lines). A: Waveforms recorded at electrodes contralateral and ipsilateral to the target. B: Difference waveforms obtained by subtracting ipsilateral from contralateral ERPs. The CDA time window (400-800 ms) is marked in grey. p -values reflect the comparison between the two-items and single-item difference waves in the early 400-500 ms window and the later 500-800 ms window.

N2pc components: Correct first report versus distractor intrusion first report. Figure 8A shows the ERP waveforms triggered by the target frame at electrodes PO7 and PO8 contralateral and ipsilateral to the target in the 500 ms interval after target frame onset, separately for trials where the first response was correct or a distractor intrusion. The corresponding difference waves obtained by subtracting ipsilateral from contralateral ERPs are shown in Figure 8B. Clear N2pc components were present for both types of trials, but there was a marked N2pc onset latency difference, with an N2pc delay on trials where distractor intrusions were reported. This was

confirmed by the analysis of N2pc onset latencies, based on a 50% average peak amplitude criterion ($M = -0.9 \mu\text{V}$). The N2pc component emerged 19 ms earlier on trials with correct responses relative to distractor intrusion trials, $M = 207.2 \text{ ms}$ vs. $M = 226.1 \text{ ms}$, and this difference was significant, $F_{adjusted}(1,14) = 6.99$, $p = .02$. N2pc mean amplitudes measured in the 200–300 ms time window were significantly different from zero both on trials with correct responses and on distractor intrusions trials, both $ps < .001$. The mean N2pc amplitude was numerically larger on trials where the target was reported correctly relative to intrusion trials ($M = -1.60 \mu\text{V}$ vs. $M = -1.37 \mu\text{V}$), but this difference was not statistically significant, $F(1,14) = 2.26$, $p = .16$, $\eta_p^2 = .14$.

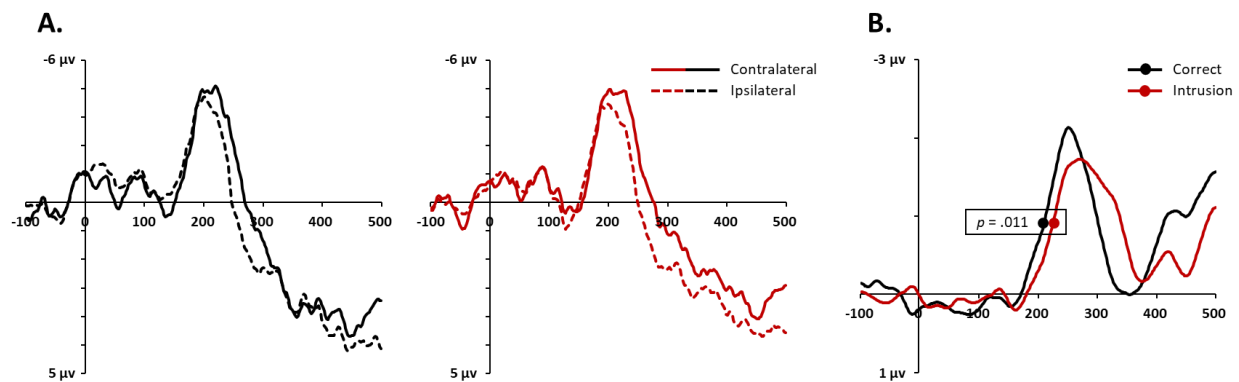


Figure 8. Grand-average event-related potentials (ERPs) waveforms on electrodes PO7/PO8 elicited in Experiment 4 by target frames, shown separately for trials where participants reported the target on the first response (correct, black lines) and trials where participants reported the PTD (intrusion, red lines). A) Waveforms recorded at electrodes contralateral and ipsilateral to the target. B) N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs. N2pc onset latencies are indicated by dots. In line with the N2pc onset analyses, a 10 Hz low-pass filter was applied to these waveforms.

Exploratory N2pc analysis: Single-item versus two-items reports. As can be seen from Figure 7B, the N2pc appeared to emerge later on single-item relative to two-items report trials. Although we had no a priori predictions regarding the presence of such an N2pc onset latency difference, we assessed its reliability on the basis of the EEG epochs used in the CDA analysis. After applying a

10-hz lowpass filter to these data, we calculated N2pc onset latencies using the procedures described above, based on a 50% average N2pc peak amplitude criterion ($M = -1.13 \mu\text{V}$). This analysis revealed a significantly earlier N2pc onset on two-items report trials ($M = 217 \text{ ms}$) relative to single-item report trials ($M = 230 \text{ ms}$), $F_{\text{adjusted}}(1,14) = 5.21, p = .04^5$.

Discussion

Experiment 4 yielded three clear-cut results. Confirming the findings of Experiment 1, the likelihood that participant reported the target after first reporting the PTD, and the probability of reporting the PTD after a first correct response were both higher than chance. This indicates that both items were encoded in WM on at least a subset of trials. Second, the N2pc emerged earlier when the first response was correct than when participants chose the PTD for their first report. This confirms our previous observations (Zivony & Eimer, in press), and shows that distractor intrusion errors are linked to a delay in the onset of attentional engagement. Third, and most relevant for the issues addressed in Experiment 4, there were reliable CDA amplitude differences between trials where participants reported only a single item (either the target or the PTD) and trials where both items were reported. CDA components were smaller on single-item as compared to two-items report trials, strongly suggesting that these two types of trials differed in the number of items that were held in WM. Importantly, this difference was already present at the time when the CDA emerged, and there was no indication that it increased during the later phase of the retention interval. These observations provide no support for a catch-and-release account

⁵ In this analysis, the single-item report N2pc waveforms included more trials with intrusion responses than correct responses. Similarly, the two-items report N2pcs included more intrusion-first trials than correct-first trials. Because these unequal trial numbers may have affected estimated N2pc onsets, we repeated the same analysis while giving each response condition an equal weight in the calculation of N2pc difference waves for both types of trials. The N2pc onset latency delay for single-item as compared to two-item report trials was still present (230 ms versus 217 ms), although this difference was now only marginally significant, $F_{\text{adjusted}}(1,14) = 3.85, p = .07$.

according to which both items are encoded but one of them is then dropped from WM on single-item report trials. They are however in line with the hypothesis that competition at the encoding stage prevents either the target or the PTD from entering WM on a substantial proportion of trials.

It is notable that the CDA was already reliably present from 400-500 ms after target onset only when both items were reported, and emerged only after 500 ms on single-item report trials, indicative of a delay of WM encoding on these trials. An analogous onset difference was also observed for N2pc components. The N2pc was delayed on single-item as compared to two-items report trials, similar to the delay observed for trials where participants picked the PTD rather than the target for their first report (see above). These observations suggest that slower attentional engagement has a knock-on effect on the speed of WM encoding, and reduces the probability that both items will be encoded. This would be in line with previous work on interactions between the efficiency of attentional and WM processes in single-frame displays (e.g., Adam et al., 2015; Salahub et al., 2019). The results of Experiment 4 suggest that trial-by-trial variability in the speed of attentional engagement in RSVP streams can affect the competition for entry into WM between items presented in rapid succession. Inhibitory interactions at a perceptual stage should have more pronounced effects the more time passes before attention is engaged. Therefore, on trials where attentional engagement is delayed, it is more likely that at least one representation would become too degraded to gain access to WM.

Even though Experiments 1 and 4 employed similar procedures, the percentage of trials where participants reported both the target and the PTD was higher in Experiment 4 relative to Experiment 1 (61% versus 35%; see also Figures 2B and 7B). One possible reason for this difference is that Experiment 4 included twice as many trials as Experiment 1, and that trials where

both these items are reported become more frequent after extended task practice. To test this, we compared the frequency of two-item report trials in the first six blocks and in the final six blocks of Experiment 4. This comparison revealed that two-items reports were more frequent in the last half of the experiment relative to the first, $M = 58.4\%$ vs. $M = 63.9\%$, $t(14) = 2.76$, $p = .02$, $d = 0.71$, compatible with a practice effect. However, another and more interesting possibility is that this difference chiefly reflects a difference in task strategies between these two experiments. In contrast to Experiment 1, the target in Experiment 4 was always followed by a category-matching PTD, and this could have encouraged participants to select and encode two items, in a manner similar to attentional blink tasks with two successively presented targets. Indeed, a comparison between the two experiments revealed a significant difference, even when we excluded the second half of Experiment 4 from the analysis, $t(29) = 3.72$, $p = .002$, $d = 1.38$. We discuss this possibility further in the General Discussion.

General Discussion

When a target in an RSVP stream task is followed by a distractor that shares its response dimension, participants often erroneously report the post-target distractor (PTD) instead of the target. Such distractor intrusions have been demonstrated in numerous studies that employed a wide array of tasks (e.g., identity reports: Goodbourn et al., 2016; color reports: Gathercole & Broadbent, 1984; shape reports: Livesey & Harris, 2011), presentation rates (Kikuchi, 1996; Lawrence, 1971) and presentation configurations (e.g., single stream: Vul et al., 2009; Recht et al., 2019; multiple streams: Goodbourn & Holcombe, 2015; Zivony & Lamy, 2016). These effects suggest that the target and distractor compete for access to perceptual report, but it is not clear whether this competition takes place prior to WM encoding, or at later stage, after both items have

been encoded.

If these items compete only after both have been encoded, it follows that on trials where one of these items is reported (indicating that an attentional episode has been successfully triggered), participants should also be able to report the second item, given the opportunity to do so. Our results provided clear evidence against this post-encoding competition hypothesis. Accuracy was consistently reduced on trials with a category-matching PTD relative to baseline trials where the target digit was followed by a letter. This was the case even when participants could provide two perceptual reports (Experiment 1) and when the post-target distractor was not a possible response alternative (Experiments 2 and 3). These observations suggest that the competition between targets and PTDs increased the likelihood that one of them would be excluded from access to WM. However, they could also be accounted for by competitive interactions within WM, which might lower accuracy by restricting the target's access to visual awareness (Pincham et al., 2016) or reducing the quality of its representation in WM. Our results provided evidence against both alternatives. Experiment 2 demonstrated that on trials where a potentially intruding PTD was present but unavailable for response, low confidence judgments accurately reflected participants' inability to select the target. If targets had been blocked from conscious report but not from encoding on these trials, a dissociation between confidence and accuracy should have been found. In Experiment 3, the probability that participants reported an item that was physically similar to the target was unaffected by the presence or absence of a potentially intruding PTD. If competitive interactions in WM had impaired the precision of target representations in WM, these errors should have been more frequent on trials where this PTD was present. Finally, Experiment 4 provided converging electrophysiological evidence for the encoding competition account. CDA amplitudes measured prior to perceptual reports were larger on trials where participants reported

both the target and the PTD relative to when they reported only one of these items. As CDA amplitudes are a marker of the number of items that are stored in WM (Vogel & Machizawa, 2004), this finding provides direct on-line evidence that the number of items encoded in WM was on average lower on trials with single item reports, indicating that only a single item was encoded on a substantial number of trials. Importantly, the CDA amplitude difference between single-item and two-items report trials was present from the moment when CDA components emerged, which is incompatible with the possibility that both items are initially encoded but one of them was later dropped from WM. Overall, these CDA results demonstrate that the competition between the target and the PTD frequently results in one of them being blocked from entering WM.

While our results are incompatible with post-encoding competition accounts, they are also not in line with the encoding competition account that we originally proposed (Zivony & Eimer, in press). According to this account, the competition between the target and the PTD at a perceptual stage results in only one of these items being encoded, either the target on trials with correct responses or the PTD on trials with intrusion errors. This version of our account was clearly refuted in the present study. The behavioural results from all four experiments demonstrated that reports of the second of these two items exceeded guess rates, suggesting that both could be represented in WM at the same time. This was confirmed with ERP markers in Experiment 4, where the increase of CDA amplitudes on two-items are compared to one-item report trials demonstrated that the number of items maintained in WM differed between these trials.

Given these results, we propose a modified version of our original encoding competition account. This modified account stipulates that the perceptual competition between the target and PTD does not always block one item from gaining access to WM, but instead reduces the likelihood that one of these items will be encoded. Whether one or both items are encoded depends

on whether the activation levels of their perceptual representations cross an *encoding threshold* (Wyble et al., 2009; 2011). When both representations are sufficiently activated, both enter WM. However, on a large subset of trials, only one representation is sufficiently activated to cross the encoding threshold. On these trials, this representation gains access to WM, whereas the other remains unstable and quickly dissipates, and therefore cannot be selected for perceptual reports. Importantly, these trial-by-trial fluctuations in activation levels depends on intertrial variability in the speed of attentional engagement processes (see Zivony & Eimer, in press), which can bias the perceptual competition between items in favour of either the target or the PTD representation. Following engagement, the activation of all representations is transiently enhanced within a brief attentional window. When engagement is triggered late, this enhancement will mainly facilitate the representation of the PTD, whereas the sensory representation of the target often remains below the encoding threshold, resulting in intrusion errors. In contrast, fast engagement boosts the activation of target representations, thus increasing the probability that the target is encoded and correctly reported. The fact that N2pc components emerged earlier in Experiment 4 on trials with correct responses relative to distractor intrusion trials (confirming the results of our previous study; Zivony & Eimer, in press) provides on-line ERP evidence for this link between attentional engagement speed and report accuracy. It is notable that in Experiment 4, N2pcs were also triggered more rapidly when both items were reported relative to single-item report trials. This suggests that when attentional engagement is fast, the representation of the PTD can be sufficiently strong to cross the encoding threshold. The hypothesis that slow attentional engagement specifically reduces the probability that target representations reach the encoding threshold, while both items are often encoded when engagement is fast is in line with the fact that on two-items report trials in Experiment 4, the PTD was more often picked for the second report than the target

(although this difference was not significant in Experiment 1).

This modified encoding competition account can explain why the target was blocked from entering WM on a proportion of the trials where the PTD was encoded, as was observed in all four experiments of the current study. This observation has important implications for models of temporal attention. First, it contradicts the widely held assumption that selection of a single pre-defined target in RSVP streams is an undemanding task (Chun & Potter, 1995; Di Lollo et al., 2005; Jolicœur & Dell'Acqua, 1998; Taatgen et al., 2009). Moreover, it is incompatible with previous models of temporal attention suggesting that all items that appear during the same attentional episode are usually encoded in WM (Olivers & Meeter, 2008; Wyble et al., 2011). These models concede that the encoding of multiple successive items within a single attentional episode can result in some costs, such as loss of temporal acuity (Hilkenmeier et al., 2012; Olivers, et al., 2011) or reduced visual awareness and confidence about the targets' presence (Pincham et al., 2016). Instead, we conclude that competitive interactions between a target and a category-matching PTD often take place at an earlier pre-encoding stage, and can have the more severe effect of preventing the target from being encoded in WM. This conclusion is however consistent with another account of attentional episodes in visual perception (eSTST model; Wyble et al., 2009, 2011), which assumes that perceptual interference between the processing of multiple successive items within a single attentional episode can reduce the likelihood that these items are successfully encoded. However, while the eSTST model suggests that this interference should be relatively weak in cases where it affects only two successive items (e.g., Dell'Acqua et al., 2012), the results of Experiments 1-3 demonstrated that the presence of a category-matching distractor can produce robust interference with target encoding processes, as reflected by substantial performance costs relative to baseline trials.

While our findings may challenge the generality of models that propose a post-encoding locus of competition, it is important to note that they do not invalidate them, as these models were specifically designed to describe conditions such as attentional blink tasks where participants search and report *two targets*. In these cases, both targets are usually reported with high accuracy when they appear successively in the same location. This is true even in tasks with multiple streams similar to the design used here (e.g., Jefferies et al., 2015, 2019; Tan & Wyble, 2015; Verleger et al., 2011). In contrast, participants in the current study were specifically instructed to find and report a *single target*. This difference in task instructions may be theoretically important, as it is possible that participants might adopt different top-down control settings when searching for a single as compared to multiple targets. We suggest that there should be considerable flexibility to calibration encoding thresholds in line with current task demands (see Figure 9). A fixed encoding threshold may not be adaptive, given that the visual system has to detect and encode task-relevant events in a wide variety of different spatiotemporal contexts. In particular, given that successive items in RSVP streams compete with each other, lowering encoding thresholds in situations where more than one target is expected is strategically sensible, because it would allow multiple relatively weakly activated item representations to be encoded. In this case, despite the perceptual competition between items that appear inside the attentional episodes, more than one representation is likely to cross the threshold and gain access to WM. Initial indirect evidence for this possibility comes from a comparison of the results of Experiments 1 and 4. In Experiment 4, where the target was never followed by a category-nonmatching distractor (i.e., there were no baseline trials), participants were much more likely to report both the target and the PTD than in Experiment 1, where baseline trials were included. It is possible that some participants in Experiment 4 had noticed that two candidate target items (i.e., digits) always appeared in

immediate succession in the same stream, and had lowered their encoding threshold in order to increase the probability that both of these items would be encoded. Figure 9 represent an illustrative example of how such changes would affect encoding and subsequent perceptual reports. This hypothesis require testing in future studies that directly compare encoding under conditions where participants expect one target versus multiple targets.

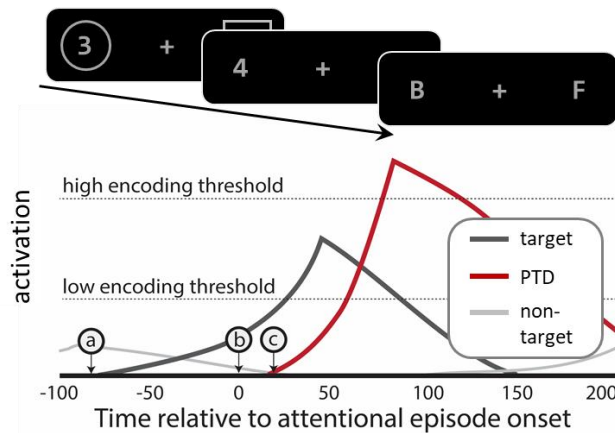


Figure 9. Illustration of the relationship between an encoding threshold and the number of encoded items in a distractor intrusion task. The x-axis reflects time relative to the onset of the attentional episode and the y-axis reflects the strength of the perceptual representation. The curves reflect the activation of the target (dark grey), post-target distractor (red) and other non-targets (light grey). In this example, the attentional episode is triggered (reflected by the letter b) relatively late, 80 ms after the visual input from the target reaches the visual cortex (a) and 20 ms before the visual input from the post-target distractor (PTD) reaches the visual cortex (c). Accordingly, activation of the PTD representation will be more strongly enhanced than the target representation. Objects are encoded if they cross an engagement threshold (dotted line). If participants expect only one target, they will set their encoding threshold at a high point, resulting in encoding of only one item on many trials. If they expect two targets, they will set their encoding threshold at a lower point, allowing both items to be encoded. In both cases non-targets that are outside the attentional episode and category-nonmatching items will not cross the encoding threshold.

The hypothesis that the encoding threshold can be adjusted in line with task demands may also be able to explain findings from the broader temporal attention literature, and generate novel testable predictions. For example, some individuals are not affected by the attentional blink, and

these participants show a faster P3 latency (Martens et al. 2006), a component that is often taken as an index of WM updating (Polich, 2007). These “non-blinkers” may have a lower encoding threshold, which would allow them to successfully encode both successive target objects, even when the second target is only weakly activated. It would be interesting to study whether blinkers and non-blinkers also differ with respect to the frequency of PTD distractor intrusion errors, or their ability to report both the target and the PTD in two-report tasks. Notably, the hypothesis that encoding thresholds are flexible also yields the counterintuitive prediction that instructing participants to search for a single target should result in lower accuracy relative to a task where they have to search for two successive targets. If they adopt a higher encoding threshold in the single target search, this will frequently result in only the post-target distractor being encoded, whereas a lower threshold during two-target search will enable them to encode both targets on the majority of trials.

Finally, our modified encoding competition account may help explain results from the object substitution masking (OSM) paradigm, where a target is surrounded by four dots, and identification accuracy is impaired when the offset of the dots is delayed relative to target offset (Enns & Di Lollo, 1997). Since the target and the dots are processed in the same attentional episode, the ensuing perceptual competition can block the target from being encoded in WM, in particular when the dots stay present after target offset and therefore remain more strongly activated. Thus, correct target reports in OSM tasks might be observed when both the target and the mask are encoded in WM, a conclusion that is consistent with the finding that CDA components are initially larger on delayed-offset trials with correct as compared to incorrect reports (Salahub & Emrich, 2018).

Conclusion

Competition between targets and distractors that appear in RSVP streams during the same attentional episode can result in distractor intrusions. Using a combination of behavioural and electrophysiological measures, we demonstrated that this competition occurs at a perceptual stage and reduces the likelihood that the target gains access to WM. This suggests that attentional engagement is not a sufficient condition for objects that appear within a single attentional episode to be encoded in WM. We conclude that only items whose activation level crosses a threshold level will be encoded. Trial-by-trial fluctuations in the speed of attentional engagement bias perceptual competition between successive items, and thus the probability that one or several of them will enter WM. We also speculate that encoding thresholds can be adjusted based on task demands. This account can explain a wide variety of findings in the temporal attention literature and chart new paths of research into the relationship between temporal aspects of selective attention and WM encoding.

Acknowledgments

This work was supported by a Newton grant from the British Academy (grant number NIF\R1\180384) to A. Zivony. The data for all behavioural analyses and for the CDA and N2pc analyses are posted at <https://doi.org/10.6084/m9.figshare.12464870>.

References

Adam, K. C., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The contribution of attentional lapses to individual differences in visual working memory capacity. *Journal of Cognitive Neuroscience*, 27(8), 1601-1616.

- Avital-Cohen, R., & Tsal, Y. (2016). Top-down processes override bottom-up interference in the flanker task. *Psychological Science*, *27*(5), 651-658.
- Balaban, H., Drew, T., & Luria, R. (2018). Delineating resetting and updating in visual working memory based on the object-to-representation correspondence. *Neuropsychologia*, *113*, 85-94.
- Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, *55*(4), 412-428.
- Berggren, N., & Eimer, M. (2016). Does contralateral delay activity reflect working memory storage or the current focus of spatial attention within visual working memory?. *Journal of Cognitive Neuroscience*, *28*(12), 2003-2020.
- Berggren, N., & Eimer, M. (2019). The roles of relevance and expectation for the control of attention in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(9), 1191-1205.
- Botella, J., & Eriksen, C. W. (1992). Filtering versus parallel processing in RSVP tasks. *Perception & Psychophysics*, *51*(4), 334-343.
- Botella, J., Suero, M., & Barriopedro, M. I. (2001). A model of the formation of illusory conjunctions in the time domain. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1452-1467.
- Brisson, B., Robitaille, N., & Jolicœur, P. (2007). Stimulus intensity affects the latency but not the amplitude of the N2pc. *Neuroreport*, *18*(15), 1627-1630.
- Broadbent, D. E. (1982). Task combination and selective intake of information. *Acta Psychologica*, *50*(3), 253-290.
- Chanceaux, M., Mathôt, S., & Grainger, J. (2014). Effects of number, complexity, and familiarity of flankers on crowded letter identification. *Journal of Vision*, *14*(6), 7-7.

- Chun, M. M. (1997). Temporal binding errors are redistributed by the attentional blink. *Perception & Psychophysics*, *59*, 1191–1199.
- Chun, M. M. & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109–127.
- Dell'Acqua, R., Dux, P. E., Wyble, B., & Jolicoeur, P. (2012). Sparing from the attentional blink is not spared from structural limitations. *Psychonomic Bulletin & Review*, *19*(2), 232-238.
- Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, *29*(5), 761-778.
- Di Lollo, V., Kawahara, J., Shahab Ghorashi, S. M., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research*, *69*, 191–200.
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225-234.
- Eimer, M., Kiss, M., Press, C., & Sauter, D. (2009). The roles of feature-specific task set and bottom-up salience in attentional capture: an ERP study. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1316-1328.
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological science*, *8*(2), 135-139.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2013). G*Power (Version 3.1.7) [Computer software]. Kiel, Germany: University of Kiel.
- Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics*, *64*(5), 741-753.

- Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications*, *3*(1), 1-8.
- Gathercole, S. E., & Broadbent, D. E. (1984). Combining attributes in specified and categorized target search: Further evidence for strategic differences. *Memory & Cognition*, *12*(4), 329-337.
- Goodbourn, P. T., & Holcombe, A. O. (2015). "Pseudoextinction": Asymmetries in simultaneous attentional selection. *Journal of experimental psychology: human perception and performance*, *41*(2), 364-384.
- Goodbourn, P. T., Martini, P., Barnett-Cowan, M., Harris, I. M., Livesey, E. J., & Holcombe, A. O. (2016). Reconsidering temporal selection in the attentional blink. *Psychological Science*, *27*(8), 1146-1156.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: a Brunswikian theory of confidence. *Psychological Review*, *98*(4), 506-528.
- Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-color and multiple-color visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 86-101.
- Grubert, A., Krummenacher, J., & Eimer, M. (2011). Redundancy gains in pop-out visual search are determined by top-down task set: Behavioral and electrophysiological evidence. *Journal of Vision*, *11*(14), 1-10.
- Guggenmos, M., Wilbertz, G., Hebart, M. N., & Sterzer, P. (2016). Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *Elife*, *5*, e13388.
- Harvey, N. (1997). Confidence in judgment. *Trends in Cognitive Sciences*, *1*(2), 78-82.

- Hilkenmeier, F., Olivers, C. N., & Scharlau, I. (2012). Prior entry and temporal attention: Cueing affects order errors in RSVP. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 180-190.
- Hommel, B., & Akyürek, E. G. (2005). Lag-1 sparing in the attentional blink: Benefits and costs of integrating two events into a single episode. *The Quarterly Journal of Experimental Psychology Section A*, *58*(8), 1415-1433.
- Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of Neurophysiology*, *103*(4), 1963-1968.
- Intraub, H. (1985). Visual dissociation: An illusory conjunction of pictures and forms. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 431-442.
- Jachs, B., Blanco, M. J., Grantham-Hill, S., & Soto, D. (2015). On the independence of visual awareness and metacognition: A signal detection theoretic analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 269-276.
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). New York, NY: Oxford University Press.
- Jefferies, L. N., Roggeveen, A. B., Enns, J. T., Bennett, P. J., Sekuler, A. B., & Di Lollo, V. (2015). On the time course of attentional focusing in older adults. *Psychological research*, *79*(1), 28-41.
- Jefferies, L. N., Enns, J. T., & Di Lollo, V. (2019). The exogenous and endogenous control of attentional focusing. *Psychological research*, *83*(5), 989-1006.
- Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, *36*, 138-202.

- Jonides, J., & Gleitman, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Perception & Psychophysics*, *12*(6), 457-460.
- Kikuchi, T. (1996). Detection of Kanji words in a rapid serial visual presentation task. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(2), 332-341.
- Kiss, M., Van Velzen, J., & Eimer, M. (2008). The N2pc component and its links to attention shifts and spatially selective visual processing. *Psychophysiology*, *45*(2), 240-249.
- Lawrence, D. H. (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception & Psychophysics*, *10*, 85-89.
- Livesey, E. J., & Harris, I. M. (2011). Target sparing effects in the attentional blink depend on type of stimulus. *Attention, Perception, & Psychophysics*, *73*(7), 2104.
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391-400.
- Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100-108.
- Mamassian, P. (2016). Visual confidence. *Annual Review of Vision Science*, *2*, 459-481.
- Martens, S., Munneke, J., Smid, H., & Johnson, A. (2006). Quick minds don't blink: Electrophysiological correlates of individual differences in attentional selection. *Journal of Cognitive Neuroscience*, *18*(9), 1423-1438.
- McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex*, *43*(1), 77-94.
- Miller, J., & Bauer, D. W. (1981). Visual similarity and discrimination demands. *Journal of Experimental Psychology: General*, *110*(1), 39-55.

- Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, *35*(1), 99-115.
- Olivers, C. N., & Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychological Review*, *115*(4), 836-863.
- Olivers, C. N., Hilkenmeier, F., & Scharlau, I. (2011). Prior entry explains order reversals in the attentional blink. *Attention, Perception, & Psychophysics*, *73*(1), 53-67.
- Olivers, C. N., Van Der Stigchel, S., & Hulleman, J. (2007). Spreading the sparing: Against a limited-capacity account of the attentional blink. *Psychological research*, *71*(2), 126-139.
- Overgaard, M., Timmermans, B., Sandberg, K., & Cleeremans, A. (2010). Optimizing subjective measures of consciousness. *Consciousness and Cognition*, *19*(2), 682-684.
- Pincham, H. L., Bowman, H., & Szucs, D. (2016). The experiential blink: Mapping the cost of working memory encoding onto conscious perception in the attentional blink. *Cortex*, *81*, 35-49.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128-2148.
- Popple, A. V., & Levi, D. M. (2007). Attentional blinks as errors in temporal binding. *Vision Research*, *47*(23), 2973-2981.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849-860.
- Recht, S., Mamassian, P., & de Gardelle, V. (2019). Temporal attention causes systematic biases in visual confidence. *Scientific Reports*, *9*(1), 1-9.

- Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review*, *93*(2), 180-206.
- Salahub, C. M., & Emrich, S. M. (2018). ERP evidence for temporal independence of set size and object updating in object substitution masking. *Attention, Perception, & Psychophysics*, *80*(2), 387-401.
- Salahub, C., Lockhart, H. A., Dube, B., Al-Aidroos, N., & Emrich, S. M. (2019). Electrophysiological correlates of the flexible allocation of visual working memory resources. *Scientific Reports*, *9*(1), 1-11.
- Shih, S. I. (2008). The attention cascade model and attentional blink. *Cognitive Psychology*, *56*, 210-236.
- Shih, S. I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review*, *109*(2), 260-305.
- Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, *102*, 503–532.
- Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009). Too much control can hurt: A threaded cognition model of the attentional blink. *Cognitive Psychology*, *59*, 1–29.
- Tan, M., & Wyble, B. (2015). Understanding how visual attention locks on to a location: Toward a computational model of the N 2pc component. *Psychophysiology*, *52*(2), 199-213.
- Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, *9*(1), 40-50.
- Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology*, *38*(5), 816-827.

- Verleger, R., Śmigasiewicz, K., & Möller, F. (2011). Mechanisms underlying the left visual-field advantage in the dual stream RSVP task: Evidence from N2pc, P3, and distractor-evoked VEPs. *Psychophysiology*, *48*(8), 1096-1106.
- Visser, T. A., Zuvic, S. M., Bischof, W. F., & Di Lollo, V. (1999). The attentional blink with targets in different spatial locations. *Psychonomic Bulletin & Review*, *6*(3), 432-436.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748-751.
- Vul, E., Hanus, D., & Kanwisher, N. (2009). Attention as inference: selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General*, *138*(4), 546-560.
- Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E. J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H., Derks, K., Rouder, J. N. & Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58-76.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it?. *Nature Reviews Neuroscience*, *5*(6), 495-501.
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 0058.
- Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, *400*(6747), 867-869.

- Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of experimental psychology: Human perception and performance*, 35(3), 787-807.
- Wyble, B., Potter, M. C., Bowman, H., & Nieuwenstein, M. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology: General*, 140(3), 488-505.
- Zivony, A., Allon, A. S., Luria, R. & Lamy, D. (2018). Dissociating between the N2pc and attentional shifting: an attentional blink study. *Neuropsychologia*, 121, 153-163.
- Zivony, A. & Eimer, M. (in press). Distractor intrusions are the result of delayed attentional engagement: A new temporal variability account of attentional selectivity in dynamic visual tasks. *Journal of Experimental Psychology: General*.
- Zivony, A., & Lamy, D. (2016). Attentional capture and engagement during the attentional blink: A “camera” metaphor of attention. *Journal of Experimental Psychology: Human Perception and Performance*, 42(11), 1886–1902.