Tay, D., Jannati, A., Green, J. J., & McDonald, J. J. (in press). Dynamic inhibitory control prevents salience-driven capture of visual attention. *Journal of Experimental Psychology: Human Perception and Performance.*

Dynamic inhibitory control prevents salience-driven capture of visual attention

Daniel Tay[1*], Ali Jannati[1], Jessica J. Green[2], and John J. McDonald[1*]

[1]Department of Psychology, Simon Fraser University

[2]Department of Psychology, McCausland Center for Brain Imaging, and Institute for Mind and Brain, University of South Carolina

Author Note

Ali Jannati is now at the Department of Neurology, Boston Children's Hospital, Harvard Medical School.

D.T, J.J.G., A.J., and J.J.M designed research; D.T. performed research; D.T. analyzed data; D.T. and J.J.M. wrote paper.

The authors declare no competing financial interests.

This study is supported by grants from the Natural Sciences and Research Council of Canada, Canadian Foundation for Innovation, and the Canada Research Chairs program. We thank Jennifer Hoffmeister for assistance with data collection.

Correspondence concerning this article should be addressed to Daniel Tay, Department of psychology, 8888 University Drive, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. Email: daniel_tay@sfu.ca

Data and materials are available from the authors upon request.

Abstract

The salience-driven selection theory is comprised of three main tenets: (a) the most salient stimulus within a monitored region of the visual field captures attention, (b) the only way to prevent salience-driven distraction is by narrowly focusing attention elsewhere, and (c) all other goal-driven processes are possible only after the most salient item has been attended. Evidence for and against this theory has been provided from two experimental paradigms. Here, event-related potentials (ERPs) recorded in a novel Go/No-Go paradigm disconfirmed all three of tenets of the theory. Participants were instructed to search cyan-item displays for a salient orientation singleton (Go trials) and to ignore randomly intermixed yellow-item displays that could also contain an orientation singleton (No-Go trials). ERP components associated with attentional orienting (N2pc), distractor suppression (P_D), and stimulus relevance (P2a) were isolated to test predictions stemming from the salience-driven selection theory. On No-Go trials, the salient oddball elicited a P_D rather than an N2pc, indicating that it was suppressed, not attended. Moreover, a P2a emerged before the N2pc on Go trials, demonstrating that observers first evaluated the global colour of each display and then decided to search for the oddball (Go trials) or to ignore it (No-Go trials). We conclude that goal-driven processes can lead to the prevention of salience-driven attention capture by salient visual objects within the attentional window.

Key words: distraction, attention capture, automaticity, suppression, event-related potentials

Significance Statement

It is important to understand how humans mitigate distraction to prevent injury and to optimize performance and productivity. Some researchers believe it is possible to ignore potentially distracting visual stimuli, whereas others believe that salient distractors invariably capture attention. This debate has continued because most evidence for or against salience-driven distraction is open to multiple interpretations. We resolve the debate by isolating electrical brain activity associated with attentional orienting, stimulus relevance, and proactive suppression in a dynamic search task that required participants to withhold responses to a salient stimulus on half of the trials. Our participants were able to decide on the fly to attend to salient visual stimuli or to ignore them so that they did not divert attention. We conclude that salience does not determine the order of attentional selection in visual tasks.

For decades, researchers have debated the extent to which salient visual distractors capture attention. According to one theory, observers will automatically orient attention to the most salient stimulus in the visual field unless attention is already actively focused elsewhere in the visual field (Luck et al., 2021; Theeuwes, 2010). More formally, this salience-driven selection theory is comprised of three tenets. First, the most salient stimulus within a monitored region of the visual field (called the *attentional window*) will capture attention even when the stimulus is irrelevant to the task at hand (Hickey et al., 2006; Theeuwes, 1991). Second, salience-driven attention capture can be prevented only by restricting the size of the attentional window prior to the appearance of the stimulus (Belopolsky & Theeuwes, 2010; Belopolsky et al., 2007). This tenet is based on the assumption that salience is computed only within the attentional window during the initial feedforward sweep of visual processing, and thus if an observer actively focuses attention at one location, a stimulus appearing elsewhere will be unable to capture attention because its salience is unknown. Third, besides the ability to restrict the size of the attentional window, all other goal-related processes are possible only after salience-driven attention capture has occurred (Theeuwes et al., 2000). Thus, according to the theory, observers cannot prevent capture by a salient stimulus within the attentional window, and so they must try to recover from such capture once it has occurred.

Behavioural and electrophysiological evidence against the salience-driven selection theory has been reported, first from a modified cueing paradigm (e.g., Eimer & Kiss, 2008; Folk et al., 1992) and more recently in the additional singleton paradigm (Theeuwes, 1992). Much of the electrophysiological work has focused on two lateralized components of event-related potentials (ERPs) triggered by multi-item cue displays or task-relevant search displays. One of these components, the posterior contralateral N2 (N2pc), has been associated with attention selection of individual items or subsets of items that appear in cluttered fields (Luck, 2012; Luck et al., 1997; Luck & Hillyard, 1994a, 1994b; see also Eimer, 1996; Hickey et al., 2009; Mazza & Caramazza, 2011; Tay et al., 2019), whereas the other component, the distractor positivity (P_D),

has been associated with suppression of irrelevant and potentially distracting nontargets (Gaspar & McDonald, 2014; Hickey et al., 2009). Both of these components are maximal over the posterior scalp and are isolated by comparing voltages obtained at electrodes positioned contralateral or ipsilateral to some item in the display. The N2pc is elicited by task-relevant targets as well as irrelevant cues that resemble the target in some way, while the P_D is elicited by nontargets that appear concurrently with a to-be-attended target. Although the timing of each component depends on factors such as stimulus salience and task demands, the components are often seen 200–350 ms after the appearance of a multi-item display. The N2pc is believed to reflect processes associated with spatial filtering (either upweighting of the target or down-weighting of surrounding nontargets), whereas the P_D is believed to be associated with active suppression of a stimulus or its location.

Armed with these two ERP components, researchers have reported that salient-but-irrelevant cues do not typically capture attention (i.e., elicit N2pc) unless they possess a relevant feature (e.g., Eimer & Kiss, 2008) and that salient singletons that accompany a target in the same search display do not usually capture attention and often appear to be suppressed (i.e., elicit P_D; Jannati et al., 2013; Gaspar et al., 2016; Gaspar & McDonald, 2014; Gaspelin & Luck, 2018a). Such findings appear to indicate that goal-driven control processes can prevent salience-driven attention capture, but in prior studies, prevention of capture might be due to selection history rather than “top-down control” processes per se (Awh et al., 2012). Moreover, results from a recent study indicate that although distractor suppression may be possible with small set sizes (four items or fewer), salience-driven selection occurs with larger set sizes (Wang & Theeuwes, 2020; but see Stilwell & Gaspelin, in press).

Other theoretical perspectives allow for more control of attention to keep an observer’s attention engaged on task-relevant stimuli. According to the signal suppression hypothesis, for example, salience-driven distraction is prevented by selectively down-weighting the location of a salient distractor (Sawaki & Luck, 2010). This down-weighting process is hypothesized to act

early enough to prevent capture proactively, either by acting upon representations of stimulus salience directly or upon feature maps prior to salience computation (Gaspelin & Luck, 2018a). According to the contingent capture hypothesis, the feature-based template used to search for a target determines whether any other stimulus captures attention (Folk et al., 1992). Stimuli that possess a search-template feature capture attention reflexively, whereas other stimuli do not. This hypothesis does not rule out suppression as a means to prevent capture, but it is also possible that participants simply ignore stimuli that do not possess a search-template feature.

The debate over salience-driven selection has continued for decades, largely based on data from the two aforementioned paradigms: a modified cueing paradigm (Folk et al., 1994) and the additional singleton paradigms (Theeuwes, 1991, 1992). Some compelling evidence against purely salience-driven selection has been presented, but alternative interpretations have called this evidence into question (for reviews, see Luck et al., 2021; Theeuwes, 2010). Thus, data from a new paradigm would be helpful—if not necessary—for making progress in this debate. Here, a novel Go/No-Go search task was designed to test predictions stemming from the salience-driven selection theory and, to a lesser extent, the signal suppression hypothesis. The main goal was to track processing of a salient singleton when no behavioural response was required (on No-Go trials). Behavioural data from this type of paradigm would not be very informative on its own, but ERPs can provide considerable information about stimulus processing in the absence of a behavioural response. Moreover, we opted for an equal proportion of Go trials and No-Go trials, and thus the paradigm yielded high signal-to-noise ratios for our main ERP measures without an excessively long recording session. Participants were instructed to look at a central fixation point and to detect an orientation singleton (i.e., “oddball”) in displays containing cyan items (Go trials) and to withhold responses to displays containing yellow items (No-Go trials; colours were counterbalanced across participants; Figure 1). Because there was no requirement to bring attention into a narrowly focused state at the start of each trial, we assumed that the attentional window would remain wide until the singleton

was attended. Thus, according to the salience-driven selection theory, the singleton would capture attention regardless of trial type, and any inhibitory control implemented on No-Go trials would be evident only after attention is oriented to the singleton (e.g., to withhold a manual response). In contrast, we envisaged observers adopting a strategy to process the global display colour first and to orient attention to the singleton only when a response was required (Go trials).

---------- Insert Figure 1 about here ----------

To test these hypotheses, we recorded EEG during the search task and then set out to isolate the N2pc and P_D as well as other ERP components associated with different neuro-cognitive processes. As described earlier, measurement of the N2pc and P_D would enable us to determine whether the singleton is attended or suppressed, respectively. We predicted rapid attentional orienting to the singleton on Go trials, but the various hypotheses led to divergent predictions about singleton processing on No-Go trials. If salient stimuli within the attentional window capture attention automatically, the singleton would first elicit ERP activities associated with attentional selection (N2pc) and would then elicit late inhibitory control activity to prevent responding on no-go trials. If, however, individuals actively prevent salience-driven distraction by suppressing the location of salient distractors, the singleton would fail to trigger ERP activities associated with attentional selection and would instead trigger distractor-suppression activity to prevent an unnecessary diversion of attention. If observers simply ignore items that do not match the search template, then the singleton would elicit neither the N2pc nor the P_D.

In addition to the N2pc and P_D, we isolated the no-go P3, P2a, and singleton detection positivity (SDP) components. The no-go P3 is one of two main components elicited when participants must withhold a prepared response (Bokura et al., 2001; Tillman & Wiens, 2011). It differs from the more common P3b in that it is observed as the difference between ERP voltages obtained on No-Go trials and Go trials. When Go-trial ERPs are subtracted from No-Go trial ERPs, the no-go P3 is evident as a positivity that is maximal over the midline central and

fronto-central scalp. We chose to measure the no-go P3 rather than the no-go N2 because the latter might reflect conflict monitoring (Donkers & Boxtel, 2004) or attention (Tillman & Wiens, 2011) rather than response inhibition. Rather than subtracting Go activity from No-Go activity, we subtracted No-Go activity from Go activity. Thus, if present, the no-go P3 would appear as a negative deflection rather than as a positive deflection due to the arbitrary directionality of the subtraction, which was opposite to convention.

The directionality of our Go-minus-No-Go subtraction was chosen to highlight the P2a component, which was expected to be larger (more positive) on Go trials than on No-Go trials. The P2a has been observed in a variety of paradigms as an enhanced positive voltage for task-relevant stimuli (Potts, 2004). The P2a is largest at pre-frontal recording sites, starting approximately 180 ms after stimulus onset. It is not entirely clear whether the P2a reflects enhancement of relevant features, inhibition of irrelevant features, or both, but it has generally been considered to be an ERP index of relevancy processing. Here, we predicted the P2a to be larger for Go-colour displays than for No-Go-colour displays because responses were required only for Go trials. The critical question was whether a P2a result would occur before or after the most salient item (the singleton) captured attention. The salience-driven selection theory would predict that an observer would orient attention to the singleton and then determine its colour. In that case, the P2a should appear after the N2pc.

Finally, we set out to measure a recently discovered component called the singleton detection positivity (SDP; Tay et al., 2019). Tay et al. reported that while singleton-present displays and singleton-absent displays equally elicited the P3b over the midline parietal scalp (at electrode Pz), singleton-present displays elicited considerably larger positivity over more posterior regions. The difference, which was isolated by subtracting singleton-absent ERPs from singleton-present ERPs, began 200 milliseconds after stimulus onset, lasted for ~250 milliseconds, and was maximal bilaterally over the occipital scalp (at electrodes PO7 and PO8). Because of overlap with the lateralized N2pc, the SDP was larger over the ipsilateral scalp than

the contralateral scalp in the time range of the N2pc (but was bilaterally distributed before the N2pc emerged and after it dissipated). Here, we asked whether singletons would elicit the SDP on No-Go trials as well as on Go trials.

Experiment 1

Materials and Methods

The Research Ethics Board at Simon Fraser University approved the research protocol used in this study. Data and materials are available upon request.

Participants

Twenty-four students from Simon Fraser University without history of neurological disorders participated after giving informed consent. For their participation, students received either $20 or course credit as part of a departmental research participation system. All subjects reported normal or corrected-to-normal visual acuity and were tested for normal colour vision using Ishihara colour plates prior to participation. Data from two participants were excluded from further analyses because more than 25% of their trials were contaminated by ocular artifacts (rejection criterion set in advance). Of the remaining 22 participants (mean age: 22.0 years), 12 were female and two were left-handed. This final sample size was selected a priori to give us sufficient power (0.8) to detect moderately large effect sizes (d = .65; calculated using G*Power version 3.1.9.7). This effect size was based on a recent study that employed the same singleton detection task with fewer items (but no Go/No-Go decision; Tay et al., 2019).

Apparatus

All experiments were conducted in a sound-attenuated and electrically shielded chamber dimly illuminated by DC-powered LED lighting. A height-adjustable LCD monitor running at 120 Hz presented visual stimuli. Participants sat in a chair and viewed the monitor at a distance of approximately 57 cm and made their responses using a gamepad. A Windows-based computed

controlled stimulus presentation and registered participants' button presses using Presentation (Neurobehavioral Systems, Inc., Albany, CA). A custom software (Acquire) recorded EEG from a second, Windows-based computer, which housed a 64-channel A-to-D board (PCI-6071e, National Instruments, Austin, TX) that connected to an EEG amplifier system with an input impedance of 1 GΩ (SA Instruments, San Diego, CA). The stimulus-control and EEG-acquisition computers were situated outside of the testing chamber.

Stimuli and Procedure

Each stimulus display consisted of a small, white fixation cross (0.3° × 0.3°; 0.3 cd/m^2) positioned at the middle of the display and 16 cyan (0.3° × 1.0°; x = .20, y = .35, 17.5 cd/m^2) or 16 yellow lines (x = .37, y = .57, 28.0 cd/m^2) that appeared within a 11.1° × 8.3° region around fixation. The coordinates of the lines were determined randomly, with the restrictions that all displays contain eight lines on either side of fixation without crossing the horizontal or vertical meridians and that no lines connect or overlap. Singleton-absent displays contained 16 horizontal or 16 vertical lines. Singleton-present displays were identical to singleton-absent displays except one of the 16 lines was replaced with a line of an orientation orthogonal to that of the surrounding lines. The resulting eight types of displays (colour × singleton presence × orientation) were randomly intermixed and presented with equal probability. Each display was presented for 750 ms, and the time between stimulus onset varied randomly between 1,350 ms and 1,650 ms. The colour of the lines indicated whether a given trial was Go or No-Go. For half of the participants, the cyan displays were used for Go trials and the yellow displays were used for No-Go trials. The colour assignment was reversed for the remaining participants (cyan = No-Go; yellow = Go). On Go trials, participants were asked to indicate the presence or absence of the singleton by pressing either the left or right shoulder button on a gamepad using their index fingers. The stimulus-response mapping was counterbalanced across participants. On No-Go

234 trials, participants simply waited for the trial to end without providing a response. Each
235 participant completed 40 blocks of 40 trials, yielding a total of 1,600 trials.

236 ***Behavioural Analysis***

237 Median RTs for target-absent and target-present trials were computed separately for each
238 participant and then were averaged across participants. Mean RTs for target-present and target-
239 absent trials were compared using a two-tailed, paired-sample *t* test. In an exploratory analysis,
240 half of the target-present trials were then subdivided based on whether the preceding trial
241 contained a target or a distractor in the same quadrant or a different quadrant. Location-priming
242 effects (i.e., faster responses following a same-location target than a different-location target;
243 Maljkovic & Nakayama, 1996) were assessed separately following a target-present (Go) trial
244 and a distractor-present (No-Go) trial using a two-way, repeated-measures ANOVA (singleton
245 relevance $\times$ location) followed by two-tailed, paired sample *t* tests with Bonferroni correction for
246 multiple comparisons (per-test α = 0.05/2). Because singletons could appear at random
247 locations, this analysis was based on the quadrant of the singleton rather than its precise
248 location. The analyses excluded trials on which participants responded incorrectly, too quickly
249 (RT < 100 ms), or too slowly (RT > 1,350 ms).

250 ***Electrophysiological Recording and Analysis***

251 EEG signals were recorded from 25 sintered Ag/AgCl electrodes positioned at standard 10-
252 10 sites (FP1, FPz, FP2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, PO7, POz,
253 PO8, O1, Oz, O2, M1). During recording, all EEG signals were referenced to an electrode
254 positioned on the right mastoid, and the ground electrode was positioned over the midline
255 frontal scalp at site AFz. To track horizontal eye movements, an additional pair of electrodes
256 placed 1 cm lateral to the external canthus of each eye recorded horizontal electrooculographic
257 (EOG) activity. Eye blinks were monitored using the FP1 electrode and all electrode
258 impedances were kept below 15 kΩ. EEG and EOG signals were amplified with a gain of

20,000, filtered using a bandpass filter of 0.01-100 Hz (two-pole Butterworth), and digitized at 500 Hz. The EEG signals were stored on a computer for offline averaging. A semi-automated procedure was performed to remove epochs of EEG that were contaminated by horizontal eye movements, blinks, or amplifier blocking. Specifically, differences between minimum and maximum voltages on the HEOG (horizontal eye movements) and FP1 (vertical eye movements and blinks) were compared to pre-set thresholds (in DAQ units, not μV). An epoch was excluded from subsequent averaging procedures when the difference between the minimum and maximum voltages exceeded the threshold for at least one type of artifact. Thresholds were determined by visually inspecting the continuous EEG and EOG to determine values that would produce rejections of all clearly visible artifacts but of few artifact-free epochs. The minimum and maximum voltages were selected within a 700-ms time window within the recording epoch that started 200 ms before the onset of the search display. Artifact-free data were then low-pass filtered (half-power cutoff) at 30 Hz to create averaged ERP waveforms. Each EEG channel was digitally rereferenced to the average of the left and right mastoid channels. The grand-averaged event-related EOG deflections were required to be below 2 μV. Positive voltages were plotted downward by convention.

The majority of statistical tests were two-tailed. However, one-sample tests of signed areas were one-tailed because of the directional nature of the measurement. Because of the inherent difficulty in asserting null hypotheses using conventional *t* tests, we computed the JZS Bayes Factor (*BF*) for all nonsignificant parametric statistics. A default scale *r* (Cauchy scale) value of .707 was used to compute all *BF*s. We reported BF_{01} values to denote the relative likelihood of observing the data given the null hypothesis is true relative to observing the data given the alternative hypothesis is true.

Due to the novelty of this paradigm, we had little a priori knowledge of the exact timing of the ERP components of interest. To avoid cherry-picking our measurement windows based on observed data, magnitudes of all ERP components were initially quantified as the signed area

within wide time windows, at electrode locations selected in advance (based on prior studies). To determine the presence of any component of interest, a nonparametric permutations approach was used to compare the measured signed area from a grand-averaged waveform to the signed area that would occur in the complete absence of the signal (i.e., on the basis of noise alone; see Sawaki et al., 2012). This was accomplished by randomly reassigning the parameters of the trial type and re-computing grand-averaged ERPs. For example, to determine whether an N2pc was present on No-Go trials, the side of the singleton (left, right) was reassigned randomly before grand-averaging. Such reassignment removes the lateralized ERP signal to enable computation of signed area due to noise on one permutation. This process was repeated 500 times to yield 500 permutations of a grand-averaged ERP. The signed positive or signed negative areas obtained from these permutations were used to provide a distribution of values expected if a null hypothesis were true. In line with the traditional threshold for statistical significance, the observed grand-averaged ERP component would be considered statistically present if the measured signed area fell beyond the 95th percentile of the estimated noise distribution. The *p* value for this permutation test was calculated using the following equation (Phipson & Smyth, 2010; see also Gaspelin & Luck, 2018a):

$$P = \frac{1 + (number\ of\ permuted\ values\ \geq\ observed\ area)}{1 + total\ number\ of\ permutations}$$

Because the permutations approach does not produce a conventional statistic (*t* or *F* value), the signed area measurement was complemented by measurement of mean amplitudes. The mean-amplitude measurement windows were contained within the wider signed-area windows and, whenever possible, were chosen a priori based on previous research. Differences in mean amplitudes across conditions (e.g., Go vs. No-Go) or versus zero microvolts (to determine presence of a component) were then assessed using pairwise *t* tests and used to compute effect sizes (Cohen's *d*).

ERPs elicited by displays containing a singleton in the left or right visual field were combined in such a way as to produce waveforms recorded contralateral and ipsilateral to the target. To isolate the N2pc and P_D, ipsilateral ERPs were subtracted from corresponding contralateral ERPs to produce contralateral-minus-ipsilateral difference waves. N2pc measurements were taken from the contralateral-minus-ipsilateral difference waves elicited by all singletons on any given trial type (Go or No-Go). N2pc magnitude was first quantified as the signed negative area within a 200–400-ms window at electrodes PO7/8 then as mean amplitude within a 275–325-ms window at the same electrode sites. The relatively late measurement window was determined based on pilot studies from the lab that suggested a delay in N2pc onset latency with the addition of a Go/No-Go element to the task. The electrode site was chosen a priori based on many prior studies that measured N2pc from PO7 and PO8 (originally called OL and OR; e.g., Eimer, 1996). Presence of an N2pc was assessed using the permutation approach, separately for Go and No-Go trials. N2pc magnitudes on Go and No-Go trials were then compared using a paired-sample *t* test. Onset latency of the N2pc was quantified as the time at which it first reached 25% of its peak amplitude, using a standard jackknife approach (Miller et al., 1998).

P_D measurements were also obtained from the contralateral-minus-ipsilateral difference waves elicited by all singletons on any given trial type (Go or No-Go). Magnitude of the P_D was first quantified as the signed positive area within a 200–500-ms window at electrodes PO7/8 and then as the mean amplitude within a 350–450-ms window at the same electrode sites. These electrodes were selected a priori based on prior studies (e.g., Gaspar & McDonald, 2014; Hickey et al., 2009). Presence of a P_D was assessed using the permutation approach, separately for Go and No-Go trials. A paired-sample *t* test was then conducted to compare the magnitude of the P_D elicited between Go and No-Go trials.

To isolate the P2a and no-go P3, ERPs elicited by No-Go trials were subtracted from ERPs elicited by Go trials to produce Go-minus-No-Go difference waves. All measurements for the

two ERP components were taken from these difference waves. Magnitude of the P2a was first quantified as the signed positive area within a 150–350-ms window at electrode FPz then as the mean amplitude within a 180–230-ms window at the same electrode. This electrode was chosen because seminal results reported the P2a to be largest over the frontal pole (Potts et al., 1996) and because it was midline between the lateral frontal-pole electrodes used in a seminal study (Potts, 2004). We chose to use a single midline electrode to reduce the number of statistical tests and thus help control familywise error rates. Presence of the P2a was confirmed using the permutation approach. Onset latency of the P2a was measured as the time point at which the P2a first reached 25% of its peak amplitude, using the jackknife approach. This latency was then compared with that of the N2pc using a paired-sample *t* test. Magnitude of the no-go P3 was first quantified as the signed negative area within a 200–400-ms window at electrode Cz then as mean amplitude within a 250–350-ms window at the same electrode. The electrode location was selected a priori based on previous studies (e.g., Donkers & Boxtel, 2004). Presence of the no-go P3 was confirmed using the permutation approach.

Along with our theoretically motivated measurements of N2pc, P_D, and P2a, we measured a recently reported component called the singleton detection positivity (SDP; Tay et al., 2019). To isolate the SDP, singleton-absent ERPs were subtracted from singleton-present ERPs to produce present-minus-absent difference waves. All SDP measurements were taken from these difference waves. SDP magnitude was first quantified as the signed positive area then as mean amplitude at ipsilateral and contralateral PO7/8 within a 200–400-ms window. The measurement window and electrodes were chosen a priori based on the first report of SDP (Tay et al., 2019). Presence of the SDP was tested using the permutation approach, separately for Go and No-Go trials. To ensure that the SDP was not due to volume conduction from the parietally maximal P3b (i.e., the P300; Squires et al., 1975), we compared mean amplitude of the P3b elicited by singleton-present and singleton-absent trials at electrode Pz during the 200–

400-ms interval using a paired-sample *t* test. Difference in size of SDP on Go and No-Go trials was then assessed using a paired-sample *t* test.

Topographical voltage maps of ERPs were constructed by spherical spline interpolation (Perrin et al., 1989). ERPs elicited by singleton-present displays were mapped by collapsing over left and right targets and left and right electrodes such that electrodes on the left and right sides were ipsilateral and contralateral to the singleton, respectively. ERPs elicited by singleton-absent displays were mapped using the original electrode montage with left and right electrodes positioned on the left and right sides of the head, respectively. Contralateral-minus-ipsilateral difference maps were produced by first subtracting the ipsilateral topography from the contralateral topography at corresponding electrode locations, then projecting this difference topography over both sides of the head using the conventional approach (e.g., Green et al., 2008).

Results and Discussion

Roughly eight percent (8.1%) of Go trials were excluded from all analyses because responses were incorrect. Less than one percent (0.9%) of Go trials were excluded from all analyses because participants failed to respond. Another 0.3% of trials were excluded from all analyses because responses were too fast (response time, RT, < 100 ms) or too slow (RT > 1,350 ms). Less than one percent (0.2%) of No-Go trials were excluded from the ERP analyses because participants failed to withhold a manual response. Finally, 11.5% of trials were excluded because an artifact was detected in the electrophysiological recordings. Mean RTs obtained from the remaining Go trials were nearly identical for singleton-present and singleton-absent displays (610 ms and 611 ms, respectively), $t(21) = 0.35$, $p = .734$.

The results of our exploratory location-priming analyses are shown in Figure 2. An ANOVA revealed that both main effects (singleton relevance and location) were significant, Fs(1,21) ≥ 10.88, ps ≤ .003, $\eta_p^2 \geq .34$, as was their interaction, $F(1,21) = 8.03$, $p = .010$, $\eta_p^2 = .28$, BF_{10} =

7.54. In line with the usual location-priming effect (Maljkovic & Nakayama, 1996), participants were faster to respond to targets that were preceded by a same-quadrant target than a different-quadrant target (580 ms vs. 608 ms), $t(21) = 3.46$, $p = .002$, $d = 0.46$. The critical question was whether a similar effect would occur when a target-present Go trial was preceded by a distractor-present No-Go trial. Such location priming would indicate that participants attended to the location of the distractor. However, no such location-priming effect was found following a No-Go trial. That is, RTs were not statistically different after a same-quadrant distractor or a different-quadrant distractor, $t(21) = 1.66$, $p = .112$, $BF_{01} = 1.38$. This modest Bayes Factor does not provide compelling support for the null hypothesis, but the significant interaction effect provides strong support for diminished location priming by irrelevant singletons.

---------- Insert Figure 2 about here ----------

Salient Singleton Suppressed on No-Go Trials

In Figure 3a, the occipital ERP waveforms elicited by singleton-present displays are plotted separately for electrodes located contralateral to the singleton (i.e., left singleton and right electrode; right singleton and left electrode) and electrodes located ipsilateral to the singleton (left singleton and left electrode; right singleton and right electrode). Differences between the contralateral and ipsilateral waveforms are generally associated with spatially specific processing of the singleton (Luck, 2012; Woodman & Luck, 1999). Early Differences in the time range of the first positive peak (P1) have been ascribed to lateralized imbalances in sensory processing (Luck & Hillyard, 1994a), whereas later differences have been ascribed to the initial attentional selection of an item (N2pc; Luck & Hillyard, 1994b) or suppression of an item (P_D; Hickey et al., 2009), depending on the polarity of the difference. Specifically, attentional selection leads to a greater negative voltage in the contralateral waveform (the N2pc), whereas suppression leads to a greater positive voltage in the contralateral waveform (the P_D). Such differences were isolated in Experiment 1 by subtracting the ipsilaterally recorded waveform

from the contralaterally recorded waveform. In the resulting contralateral-minus-ipsilateral difference waveforms (Figure 3b), the N2pc appears as a negative potential approximately 200–300 ms after display onset and the P_D appears as a positive potential 200–450 ms after display onset, with the precise timing of each potential depending on multiple factors (Eimer et al., 2010; Gaspar & McDonald, 2014; Sawaki et al., 2012).

---------- Insert Figure 3 about here ----------

To determine whether participants attended to the singleton, the contralateral-ipsilateral difference waves were assessed for the presence of an N2pc. Unsurprisingly, an N2pc was present on Go trials (area over 200–400 ms: -45.9 μV*ms; mean amplitude over 275–325 ms: -1.1 μV), p = .008, d = 1.06), which indicates that participants attended to singletons when a detection response was required. In stark contrast, there was no clear evidence for the N2pc on No-Go trials (area: -0.6 μV*ms; mean amplitude: -0.2 μV), p = .818. An additional t test of N2pc presence on No-Go trials verified this null result, $t(21) = 1.22$, p = .238, BF_{01} = 2.33. The Bayes factor indicates that this lack of an N2pc on No-Go trials is 2.33 times more likely to be observed if the null hypothesis were true than if the alternative hypothesis were true, suggesting that participants did not orient attention to singletons when detection was unnecessary. This pattern of results is inconsistent with the salience-driven selection theory, according to which the singleton would capture attention—and thus elicit the N2pc—regardless of its task relevance (Theeuwes, 2010). The results, however, are broadly consistent with the signal suppression hypothesis as well as the contingent capture hypothesis (Folk et al., 1992; Sawaki & Luck, 2010).

To evaluate the signal suppression hypothesis more directly, we tested for the presence of a P_D in the contralateral-ipsilateral difference waves. Although our main prediction focused on the potential presence of P_D on No-Go trials, a target-elicited P_D (herein called the target positivity; P_T) often occurs after a relevant item is attended (as evidenced by an N2pc; Sawaki et al., 2012). Whereas the P_D is hypothesized to reflect proactive suppression that prevents

selection (Gaspelin & Luck, 2018a; Hickey et al., 2009, the P_T is hypothesized to reflect active termination of selective processing (Sawaki et al., 2012). Accordingly, we measured magnitudes of these positivities on both Go trials and No-Go trials. The P_D was found to be present on No-Go trials (area over 200–500 ms: 108.7 μV*ms; mean amplitude over 350–450 ms: 0.5 μV), p = .002, d = 0.54, but there was no clear evidence of the P_T on Go trials (area: 36.9 μV*ms; 0.1 μV), p = .128. This null result was verified using a t test against 0 μV, $t(21) = 0.64$, p = 527, BF_{01} = 3.73. The difference in magnitude of these positivities between No-Go and Go trials was statistically significant, $t(21) = 3.35$, p = .003, d = 0.46 (Figure 3b). The presence of P_D on No-Go trials suggests that the singleton was suppressed on such trials, thereby lending some support for the signal suppression hypothesis. However, the P_D was relatively late on No-Go trials (onset latency: 305 ms), appearing well after the onset of the target-elicited N2pc on Go trials (173 ms), $t(21) = 8.32$, $p < .001$, d = 2.01. Thus, it is unlikely that this P_D tracked a suppression process that prevented capture (and the elicitation of N2pc) proactively.

Top-Down Control Precedes Salience-Driven Attentional Selection

ERPs elicited by Go and No-Go displays were compared to determine the earliest occurrence of voluntary, “top-down” processes related to task relevance. According to salience-driven selection theory, top-down processing occurs only after the most salient item in the attentional window captures attention. To test this hypothesis, we isolated an early relevance-driven ERP component (P2a) and a later component associated with response-level inhibitory control (no-go P3) by subtracting No-Go waveforms from Go waveforms. In the resulting Go-minus-No-Go difference waves, the P2a would appear as a positive deflection and the no-go P3 would appear as a *negative* deflection (due to the directionality of the subtraction procedure). The onset latency of the N2pc (here defined as the time at which the N2pc first reached 25% of its peak amplitude) was used to track the timing of attentional selection, and the presence and timing of the P2a were used to determine whether top-down processing occurred before

selection. As shown in Figure 4, a large P2a was evident over the anterior scalp. Statistical analyses confirmed that this P2a was present at electrode FPz (area over 150–350 ms: 367.7 μV*ms; mean amplitude over 180–230 ms: 4.1 μV), $p = .002$, $d = 3.09$, and that its onset latency was significantly shorter than that of the N2pc (159 ms vs. 262 ms, respectively), $t(21) = 8.34$, $p < .001$, $d = 2.63$. These findings show clearly that processes associated with task relevance can occur "on the fly" prior to salience-driven selection, contrary to Tenet 3 of salience-driven selection theory (as formulated in the Introduction) that voluntary, "top-down" processes can only occur following salience-driven capture. A no-go P3 was also present (area over 200–400 ms: -245.2 μV*ms; mean amplitude over: 250–350 ms: -2.2 μV), $p = .002$, $d = 1.04$, which suggests that inhibitory control processes were beginning in medial frontal areas involved in performance monitoring (Smith et al., 2010).

---------- Insert Figure 4 about here ----------

Reduced Singleton Detection on No-Go Trials

In addition to testing predictions stemming from salience-driven selection theory and the competing signal-suppression hypothesis with N2pc, P_D, and P2a measurements, we performed an exploratory analysis of a recently discovered ERP component associated with detection of visual singletons (Tay et al., 2019). This component—called the SDP—was found to be maximal bilaterally over the occipital scalp. In Experiment 1, the SDP was isolated by subtracting singleton-absent waveforms from singleton-present waveforms (Figure 5a). As expected, the resulting difference waves revealed a positive potential that was largest over the occipital scalp and began approximately 200 ms after display onset (Figures 5b–d). The early onset and posterior topography differentiate the SDP from other, more common ERP positivities, such as the P3b. Unsurprisingly, the SDP was elicited by the singleton on Go trials over both the ipsilateral (area over 200–400 ms: 455.3 μV*ms; mean amplitude over 200–400 ms: 2.3 μV) and contralateral scalp (area: 417.8 μV*ms; mean amplitude: 2.1 μV), ps = .002, ds ≥ 2.05. In

contrast, P3b amplitude measured over the midline parietal scalp (Pz) did not differ across singleton-present (7.2 μV) and singleton-absent (6.7 μV) trials in the same time interval, $t(21) = 1.56$, $p = .141$, $BF_{01} = 1.57$. The Bayes factor does not offer strong support for the null hypothesis, but there was no indication that P3b amplitude was influenced by singleton presence. A small but significant SDP was observed on No-Go trials over the ipsilateral scalp (area: 88.9 μV*ms; mean amplitude: 0.4 μV) and the contralateral scalp electrodes (area: 136.6 μV*ms; mean amplitude: 0.7 μV), ps ≤ .014, ds ≥ 0.49, but this SDP was significantly smaller than that on Go trials, ts(21) ≥ 6.51, ps < .001, ds ≥ 1.50. This diminutive SDP indicates that the singleton was detected on a small proportion of No-Go trials or that a small amount of detection activity occurred on the majority of No-Go trials. Either way, the marked reduction on No-Go trials indicates that the detection process underlying the SDP was prevented or greatly attenuated on No-Go trials.

---------- Insert Figure 5 about here ----------

Experiment 2

The N2pc that was observed on Go trials of Experiment 1 emerged approximately 262 ms after onset of the singleton-present display, which is relatively late compared to most previous visual search studies. Several factors are known to influence the timing of the N2pc (e.g., Eimer et al., 2010; Gaspar & McDonald, 2014). Therefore, it is likely that an evaluation of display colour delayed the initiation of search on Go trials in Experiment 1. The task used in Experiment 1 differed from previous studies in two potentially important ways. First, the orientations of the singleton and surrounding items swapped randomly across trials, and thus participants were unable to use a feature-guided search strategy to rapidly locate the target in Experiment 1 (see also Tay et al., 2019). Second, the global colour changed unpredictably across trials in Experiment 1, whereas the colours of items are usually fixed in ERP studies of “pop out” visual search. Either one of these factors could have delayed the N2pc relative to its typical time

range. However, according to proponents of salience-driven selection theory, attention capture can be prevented only by slowing down search and deploying attention serially to inspect individual items. The presence of a target N2pc provides some evidence against this possibility in Experiment 1, since a random serial search would eliminate the N2pc entirely. Nonetheless, in Experiment 2, we set out to determine whether the Go/No-Go aspect of Experiment 1 impacted the latency of the N2pc. As in Experiment 1, the global colour of the search display changed unpredictably from blue to yellow. Here, however, participants were instructed to indicate the presence or absence of the singleton on all trials, regardless of colour (termed All-Go condition). If the Go/No-Go decision delayed the N2pc in Experiment 1, the N2pc should be found to occur earlier in Experiment 2.

Materials and Method

Participants

Twenty-five SFU students with no history of neurological disorder participated in Experiment 2 after giving informed consent. Three were omitted from the final analysis because they had excessive EEG artifacts, leaving a final sample of 22 participants (13 females; 3 left-handed; mean age: 19.0 years). All had normal or corrected-to-normal visual acuity and normal colour vision. As in Experiment 1, this final sample size was selected a priori to give us sufficient power (0.8) to detect moderately large effect sizes (d = .65).

Stimuli and Procedure

The stimuli and procedures in Experiment 2 were identical to those in Experiment 1 except participants responded to both cyan and yellow stimulus displays (i.e., All-Go condition) and the entire experiment comprised of 20 blocks of 40 trials for a total of 800 trials.

Behavioural Analysis

Median RTs for target-absent and target-present trials were computed separately for each participant. The analysis excluded trials on which participants responded incorrectly, too quickly (RT < 100 ms), or too slowly (RT > 1,350 ms). Mean RTs for target-present and target-absent trials were compared using a two-tailed, paired-sample *t* test.

Electrophysiological Recording and Analysis

The electrophysiological recording and analysis procedures were identical to those in Experiment 1, except that earlier measurement windows for signed area and mean amplitude of the N2pc were used (area: 150–350 ms; amplitude: 225–275 ms). This adjustment was made because the N2pc was expected to occur earlier when participants did not have to make a Go/No-Go decision on each trial. The measurement window used here was based on numerous other N2pc studies. In addition, N2pc onset latency obtained on Go trials of Experiment 1 and All-Go trials of Experiment 2 were compared using a two-sample *t* test. The mean-amplitude measurement window for the P_T was also shifted to be earlier (300–400 ms), on the basis that it immediately follows the N2pc. Neither P2a nor the no-go P3 were measured because No-Go trials were omitted.

Results and Discussion

Roughly eight percent (8.1%) of total trials were excluded from all analyses because responses were incorrect. Another 1.1% of trials were excluded from all analyses because participants failed to respond. Less than one percent (0.5%) of trials were excluded from all analyses because responses were too fast (RT < 100 ms) or too slow (RT > 1,350 ms). Finally, 8.5% of trials were excluded because an artifact was detected in the electrophysiological recordings.

As in Experiment 1, the mean RTs were nearly identical between target-present and target-absent trials (549 vs. 548 ms, respectively), $t(21) = 0.34$, $p = .734$. Figure 6b shows the contralateral-minus-ipsilateral difference waves from Experiment 2 (All-Go) and Experiment 1

(Go trials; replotted from Figure 2b). As expected, the N2pc was observed on singleton-present trials (area over 150–350 ms: -85.3 μV*ms; mean amplitude over 225–275 ms: -1.6 μV), p = .002, d = 1.29. Critically, the onset latency of the N2pc was 97 milliseconds shorter in Experiment 2 than on the Go trials of Experiment 1 (165 ms vs. 262 ms, respectively). This difference was found to be statistically significant, $t(42) = 2.04$, p = .048, d = 0.61, using standard jackknife procedures (Miller et al., 1998). This difference indicates that, in the Go/No-Go task, observers first evaluated the global colour of the display and then deployed attention to the singleton on Go trials and that this evaluation took roughly 100 ms, on average. Interestingly, unlike in Experiment 1, a P_T (area over 200–500 ms: 53.4 μV*ms; mean amplitude over 300–400 ms: 0.4 μV) was observed immediately following the N2pc, p = .012, d = 0.46. These results suggest that the timing of the N2pc and the presence of the P_T are affected by the complexity of a search task, with less complex tasks giving rise to an earlier N2pc and a subsequent P_T. Unsurprisingly, the SDP was observed once again, over the ipsilateral scalp (area over 200–400 ms: 497.9 μV*ms; mean amplitude over 200–400 ms: 2.5 μV) and the contralateral scalp (468.0 μV*ms; 2.3 μV), ps = .002, ds = 1.72 (Figure 6c).

---------- Insert Figure 6 about here ----------

Experiment 3

At the outset, it was assumed that observers would monitor the entire display throughout the trials of Experiment 1 in order to evaluate the global colour of each display and to search for the singleton on Go trials. This assumption was based on multiple lines of evidence showing that items are not individually selected when the task requires no item individuation (Mazza & Caramazza, 2011). However, an alternative possibility that is consistent with salience-driven selection theory (see Tenet 2 in the Introduction) is that observers adopted a strategy to restrict their attentional focus around a narrow region in the display to inspect the colour of an individual item at that location. According to proponents of the salience-driven selection theory, restricting

584 the size of the attentional window prevents attention capture because salience is computed only
585 within the window (Theeuwes, 2010). Furthermore, if attention is in a narrowly focused state at
586 the beginning of the trial (e.g., to identify colour of an item), it cannot be expanded later to
587 search for the singleton using a salience-driven strategy because salience computations within
588 the (originally narrow) window would have been driven exclusively by the initial feedforward
589 sweep of visual processing (Theeuwes, 2010). Thus, by the time the window is broadened, it
590 would be too late to update a saliency map of the display. By this account, if attention were in a
591 focused state at the outset of each trial of Experiment 1, observers would have to search for the
592 singleton by inspecting each item one-by-one until the singleton is found (a process known as
593 serial search; Treisman & Gelade, 1980). This account, however, is inconsistent with our
594 results. The presence of N2pc and P_D on Go and No-Go trials, respectively, indicate strongly
595 that the singleton was treated as the most salient item in the display. Therefore, if we accept the
596 premise that salience is computed only within the attentional window, it would follow that the
597 window must have been wide enough to encompass all items in the Go and No-Go displays.
598 Nevertheless, we conducted Experiment 3 to test this alternative, serial-search explanation
599 empirically (for more details on the explanation, see Belopolsky & Theeuwes, 2010; Belopolsky
600 et al., 2007). The stimuli and procedure of Experiment 3 were identical to those of Experiment 1,
601 except that half of the displays contained eight items instead of 16 items. If the Go/No-Go
602 procedure used in Experiment 1 causes participants to narrow their attentional windows and
603 engage in serial search, then singleton-present responses should be faster with 8-item displays
604 than with 16-item displays (for details on this approach, see Belopolsky & Theeuwes, 2010). In
605 other words, the slope of the function relating RT to the number of items in the display (set size)
606 would be *positive*, not flat (Treisman & Gelade, 1980).

607 **Material and Methods**

608 ***Participants***

609 Thirteen SFU students with no history of neurological disorder participated in Experiment 3
610 after giving informed consent. One participant was removed from the sample due to excessive
611 eye movements, leaving a final sample of 12 participants (9 females; 0 left-handed; mean age:
612 21.6 years). All had normal or corrected-to-normal visual acuity and normal colour vision. This
613 final sample size was selected a priori to give us sufficient power (0.8) to detect a large effect
614 size similar to the one reported in the focused condition of Belopolsky and Theeuwes's (2010)
615 Experiment 2 (d = 1.0).

616 ***Stimuli and Procedure***

617 The stimuli and procedures in Experiment 3 were identical to those in Experiment 1 except
618 half the stimulus arrays contained eight lines instead of 16 and the entire experiment comprised
619 of 18 blocks of 40 trials for a total of 720 trials.

620 ***Behavioural Analysis***

621 Median RTs were assessed in a two-tailed, repeated-measures ANOVA with within-subjects
622 factors for target presence (present, absent) and set size (8 items, 16 items). RTs of target-
623 present trials were then compared between 8- and 16-item displays using a paired-sample t
624 test.

625 ***Electrophysiological Recording***

626 Horizontal EOG was recorded as in Experiment 1 to monitor for eye position and reject trials
627 contaminated with eye movements. In addition, because EEG was not recorded, vertical EOG
628 was recorded bipolarly using a pair of electrodes placed above and below the left eye. The
629 vertical EOG recording was used to detect blinks and to discard trials contaminated with blinks.

630 **Results and Discussion**

631 Less than five percent (4.5%) of Go trials were excluded from all analyses because
632 responses were incorrect. None of the Go trials were excluded due to participants' failure to

633 respond. Less than one percent (0.4%) of Go trials were excluded because responses were too
634 fast (RT < 100 ms) or too slow (RT > 1,350 ms). Roughly two percent (2.2%) of No-Go trials
635 were excluded from ERP analyses because participants failed to withhold a manual response.
636 Finally, 9.9% of trials were excluded because an artifact was detected in the
637 electrophysiological recordings.

638 Figure 7 shows the RT data as functions of set size and singleton presence. An ANOVA
639 revealed that there was no main effect of singleton presence (absent: 633 ms; present: 640 ms),
640 $F(1,47) = 0.70$, $p = .422$, no main effect of set size (8 items: 638 ms; 16 items: 635 ms), $F(1,47)$
641 $= 1.19$, $p = .300$, and no interaction between the factors, $F(1,47) = 0.78$, $p = .395$. Numerically,
642 RTs for target-present trials were actually *longer* for 8-item displays than for 16-item displays
643 (643 ms vs. 636 ms, respectively). However, we did not test for a negative search slope; instead
644 we tested for a positive search slope (that is, longer RTs for the larger set size) using a one-
645 tailed test. The difference was not statistically significant, $t(11) = 1.26$, $p = .117$, $BF_{01} = 6.82$, and
646 the search slope was near zero (-0.9 ms/item). The Bayes factor indicates that the lack of a
647 positive search slope is 6.82 times more likely to be observed if the null hypothesis were true
648 than if the alternative hypothesis were true. The absence of a positive search slope disconfirms
649 the alternative, serial-search explanation for the results from Experiment 1 and indicates that
650 observers are able to prevent salience driven diversion of attention without restricting the spatial
651 extent of the attentional window. In other words, the Go/No-Go decision delayed search but did
652 not make it less efficient; participants still managed to search items in parallel as would be the
653 case without the need to make a Go/No-Go decision (Treisman et al., 1977).

654 ---------- Insert Figure 7 about here ----------

655 **General Discussion**

656 The results of the present study are inconsistent with the three main tenets of salience-
657 driven selection theory. According to this theory, the most salient item in the visual field

invariably captures attention (Tenet 1) unless an observer has already narrowed the focus of attention elsewhere in the visual field (Tenet 2). The theory also posits that besides being able to narrow their attentional focus, observers have no way to voluntarily prevent salience-driven distraction. That is, attentional orienting within the so-called attentional window is guided by stimulus salience during the initial feedforward sweep of visual processing and is only later influenced by an observer's intentions (Tenet 3). By isolating ERP activities associated with attentional selection and other processes in a novel Go/No-Go paradigm, we show that observers do not invariably orient their attention to the most salient item in the display (contrary to Tenet 1). Experiment 3 ruled out the possibility that attention was narrowly focused and serially deployed to items in search of the target singleton by showing a flat search slope that is indicative of wide attentional window (contrary to Tenet 2). Finally, ERP activity associated with an evaluation of stimulus relevance (P2a) was found to emerge before ERP activity associated with attentional selection of the most salient item in the display (target N2pc; contrary to Tenet 3). The early evaluation of stimulus relevance delayed singleton selection by approximately 100 milliseconds (Experiment 2 vs. Experiment 1).

In the present study, Go/No-Go responses were always based on global display colour and singletons were always defined on the basis of orientation. Thus, one might wonder whether our conclusions depend upon an assumption about the equality of colour and orientation processing speed. In fact, we note that we made no assumption about the equality of feature processing speeds and purposely based the Go/No-Go decision on a feature that could be processed rapidly without the need to narrowly focus attention on an individual item. According to salience-driven selection theory, every judgement about visual stimuli requires the spatial focusing of attention and such spatial focusing is determined by salience (Theeuwes, 2010). Thus, the theory predicts attention to be captured by the singleton before any item's colour can be processed, regardless of the relative processing speeds of colour and orientation. That is, the theory states that salience is computed before any specific features are selectively processed

for identification. The present results disconfirm this hypothesis regardless of the relative processing speeds of colour and orientation. We suspect that capture can be prevented even when the Go/No-Go decision takes more time (e.g., by making the Go and No-Go colours more similar), but we will leave that question for a future study.

A proponent of salience-driven selection theory might suspect that capture failed to occur on No-Go trials because the singleton was simply not salient enough to elicit capture. Several conceptual and empirical considerations argue against this possibility. On the conceptual side, salience is hypothesized to be maximal when differences between nontargets are minimized and the difference between the singleton and the nontargets is maximized (Wolfe & Horowitz, 2004). In our experiments, 15 nontarget bars all had the same orientation, and the singleton was rotated maximally from the nontargets to produce a strong saliency signal. Moreover, according to Wolfe and Horowitz, orientation is one of four visual attributes that *undoubtedly* guide attention (based on converging evidence from multiple studies). Empirically, the presence of N2pc (on Go trials of Experiment 1 and All-Go trials of Experiment 2) as well as the P_D indicate that the singleton was, in fact, treated as the most salient item in the display. Finally, and more importantly, the lack of a positive search slope in Experiment 3 provides evidence that the singleton popped out effortlessly.

Although the cognitive control processes required to inhibit a response are usually studied separately from the top-down control processes used to ignore distractors, both forms of control share two characteristics that are relevant to salience-driven selection theory. First, they are voluntary (“top-down”) processes in the sense that participants decide to make a Go/No-Go decision or try to ignore a particular stimulus. Second, both “types” of control processes would lead to a trial-by-trial evaluation of the items in each display. Critically, according to the salience-driven selection theory, neither type of voluntary process would be possible before the most salient item in the attentional window is attended. According to Theeuwes (2010), “top-down knowledge regarding non-spatial features of the objects in the visual field (such as color, shape,

luminance, etc.) cannot alter the initial selection priority" (p. 97). Therefore, the absence of N2pc on No-Go trials appears to disconfirm the salience-driven selection theory. It is possible that capture was prevented not by the explicit goals of the observer per se but by some cognitively impenetrable consequence of selection history that resulted from these goals (Awh et al., 2012; Luck et al., 2021). This possibility could be investigated in the future by providing advanced information about the need to respond on each upcoming trial and to determine whether participants can use such advance knowledge alone to ignore the singleton on No-Go trials. A genuine effect of top-down knowledge should still be found after such trial-by-trial cuing, whereas a selection-history effect should disappear.

At the outset of this investigation, we assumed that the attentional window would remain wide throughout each trial, both to identify the global colour of the display and to search for the singleton using a salience-based selection strategy. And the results of Experiments 1 and 3 support the validity of this assumption. In Experiment 1, the singleton would not have elicited an N2pc or P_D had attention been narrowly focused at the onset of the trial because salience is computed on the initial, feedforward sweep of information through the visual system only within the attentional window (Belopolsky & Theeuwes, 2010; Belopolsky et al., 2007; Itti & Koch, 2001; Theeuwes, 2010). This feedforward sweep takes upwards of 150 ms (Theeuwes, 2010), which is in line with the onset latency of the P2a component over the anterior scalp in the present study. Thus, if attention were narrowly focused to identify the colour of a single item (Go/No-Go decision), participants would have been unable to expand their attentional window and then search for the singleton using a salience-based strategy.

Our findings are generally consistent with the contingent capture hypothesis (Folk et al., 1992) and provide some evidence for the signal suppression hypothesis (Sawaki & Luck, 2010). We surmise that observers were set to process the global colour of the display at the outset of each trial, that this attentional set enabled participants to rapidly process the global colour of the display, and that the selection of a singleton was contingent upon this initial attentional set. This

provides a cogent explanation for why the singleton elicited an N2pc on Go trials but not on No-Go trials in Experiment 1. The presence of a P_D on No-Go trials is generally consistent with the signal suppression hypothesis, but the timing of this P_D is not. Namely, the No-Go P_D occurred well after the N2pc that was observed on Go trials. Thus, it is unlikely that the P_D reflected online processes associated with the proactive suppression that prevented capture (and a distractor-elicited N2pc).

However, there are at least three reasons for delayed and more sustained suppression in the current task. First, participants had to determine whether a response was required on a trial-by-trial basis. This was shown to delay the target-elicited N2pc (as evidenced by the difference in N2pc latency between Experiments 1 and 2) and may have had an even larger impact on the P_D. Second, due to the random swapping of stimuli orientation across trials, participants could not find the singleton using a feature-based template and instead had to rely on a salience-based search strategy. Prior studies have demonstrated that the target-elicited N2pc is delayed in "pure" singleton-detection tasks relative to those in feature-search tasks (Eimer et al., 2010; Tay et al., 2019), and thus distractor suppression may also be delayed in such tasks. Third, longer-lasting suppression may be required to ignore an irrelevant singleton when there is no task-relevant target in the display to occupy attention. This would result in a more sustained P_D in our singleton-detection task than in prior additional-singleton tasks involving competition between concurrently presented target and distractor singletons. Thus, while the P_D might not have tracked the initial suppression of the singleton on No-Go trials, it might have reflected suppression at later stages of processes that help observers continue to ignore the distractor.

The present findings are seemingly at odds with an attention-cueing study involving a Go/No-Go decision (Barras & Kerzel, 2016). In the prior study, each search display contained a target colour singleton that required discrimination or a nontarget colour singleton that differed in colour and required no response (No-Go trials). Each cue display also contained a colour singleton, and on any given trial, the two singletons could appear at the same location or at

762 different locations. Based on previous studies, RTs were expected to be *longer* on same-
763 location trials than on different location trials due to immediate suppression of the cue
764 (Anderson & Folk, 2012) or to cue-driven capture followed by reactive suppression (Belopolsky
765 et al., 2010). Furthermore, the RT cost was expected to be greater for cues that matched the
766 predefined nontarget than the target because of the presumed need for more inhibitory control
767 to prevent responses on No-Go trials. Barras and Kerzel did not replicate the expected
768 behavioural evidence for cue suppression or for increased inhibitory control on No-Go-colour-
769 cue trials, and they found no ERP evidence for either cue selection (N2pc) or suppression (P_D).
770 Based on these null results, the authors concluded that observers can ignore cues associated
771 with No-Go features without actively suppressing them.

772 Despite Barras and Kerzel's (2016) seemingly clear-cut conclusion, their ERPs actually
773 appear to show evidence for both the signal suppression hypothesis and the salience-driven
774 selection theory (see Barras & Kerzel, 2016, Figure 3). That is, ERPs elicited by neutral cues
775 (i.e., cue singletons possessing neither the Go colour nor No-Go colour) contained a late P_D
776 with no preceding N2pc, whereas ERPs elicited by No-Go-colour cues contained a small N2pc
777 followed by a P_T. Unfortunately, the time window used to measure the N2pc amplitude was not
778 ideal to pick out the small N2pc, and there was no measurement of the late P_D or P_T.
779 Nonetheless, neutral-colour cues may have been proactively suppressed (in line with signal
780 suppression hypothesis) and that No-Go-colour cues may have initially captured attention and
781 been suppressed reactively (in line with salience-driven selection theory). These are by no
782 means firm conclusions, and no attempt to reanalyze Barras and Kerzel's data has been
783 undertaken here. Nonetheless, the presence of a P_D in their grand-averaged waveforms
784 warrants additional research before any firm conclusions can be made about the delicate
785 balance between capture and suppression in the Go/No-Go cueing paradigm.

786 Although our predictions and conclusions were based on well-researched ERP components
787 associated with attentional selection (N2pc), signal suppression (P_D), response-level inhibition

(no-go P3), and relevance processing (P2a), we also measured a newly discovered ERP component associated with singleton detection called the SDP (Tay et al., 2019). Compared to the well-known P3b, which is largest at midline parietal sites (Squires et al., 1975), the SDP is largest over the posterior scalp. With bilateral maxima, the topography of the SDP is consistent with a pair of neural generators in left and right visual cortices. Thus, based on topography alone, the SDP and P3b appear to arise from different processes. In line with this notion, on Go trials of Experiment 1, the SDP and P3b showed different time courses as well as different topographies. Namely, the SDP reached its peak amplitude at 350–400 ms post-stimulus, whereas the P3b reached its peak amplitude at 500–600 ms post-stimulus. Notwithstanding these differences, both the SDP and P3b appeared to be attenuated for No-Go trials in the present study. The attenuation of SDP may mean that singletons went undetected on most No-Go trials or that the SDP is associated with the detection of task-relevant singletons rather than any singleton. Future research is required to shed further light on this issue.

In conclusion, the present study showed that observers are able to dynamically switch between searching for a salient visual singleton and preventing such search when the to-be-searched and the to-be-ignored displays differ on the basis of a global feature. We hypothesize that, in this situation, observers are initially set to rapidly discriminate the relevant global feature and then enter into a search mode only when it is deemed necessary. Objects are individuated only during the latter search mode, and therefore, no individual object can capture attention during the initial global-feature analysis. Critically, observers choose for themselves whether to process the global feature, and thus they are able to exert voluntary—or “top-down”—control over salience-driven distraction. Undoubtedly, such control is made easier with practice, and the selection history likely leads to an automation of the dynamic switching between searching and ignoring.

812 **References**

813 Anderson, B. A., & Folk, C. L. (2012). Dissociating location-specific inhibition and attention
814 shifts: Evidence against the disengagement account of contingent capture. *Attention,*
815 *Perception, & Psychophysics*, *74*(6), 1183-1198.
816 Awh, E., Bolopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional
817 control: a failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437-443.
818 Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception*
819 *& Psychophysics*, *55*(5), 485-496.
820 Barras, C., & Kerzel, D. (2016). Nogo stimuli do not receive more attentional suppression or
821 response inhibition than neutral stimuli: Evidence from the N2pc, PD, and N2
822 components in a spatial cueing paradigm. *Frontiers in Psychology*, *7*, 630.
823 Belopolsky, A.V., & Theeuwes, J. (2010). No capture outside the attentional window. *Vision*
824 *Research*, *50*(23), 2543-2550.
825 Belopolsky, A.V., Schreji, D., & Theeuwes, J. (2010). What is top-down about contingent
826 capture? *Attention, Perception, & Psychophysics*, *72*:326-341.
827 Belopolsky, A.V., Zwaan, L., Theeuwes, J., & Kramer, A. F. (2007) The size of an attentional
828 window modulates attentional capture by color singletons. *Psychonomic Bulletin &*
829 *Review*, *14*(5), 934-938.
830 Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response
831 inhibition in a Go/NoGo task. *Clinical Neurophysiology*, *112*(12), 2224-2232.
832 Donkers, F. C. L., & Boxtel, G. J. M. (2004). The N2 in go/no-go tasks reflects conflict
833 monitoring not response inhibition. *Brain and Cognition*, *56*(2), 165-176.
834 Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity.
835 *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225-234.

836 Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence
837 from event-related brain potentials. *Journal of Cognitive Neuroscience*, *20*(8), 1423-
838 1433.
839 Eimer, M., Kiss, M., & Cheung, T. (2010). Priming of pop-out modulates attentional target
840 selection in visual search: Behavioural and electrophysiological evidence. *Vision*
841 *Research*, *50*(14),1353-1361.
842 Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is
843 contingent on attentional control settings. *Journal of Experimental Psychology: Human*
844 *Perception and Performance*, *18*(4), 1030.
845 Gaspar, J. M., Christie, G. J., Prime, D. J., Jolicœr, P., & McDonald, J. J. (2016). Inability to
846 suppress salient distractors predicts low visual working memory capacity. *Proceedings*
847 *of the National Academy of Sciences of the United States of America*, *113*(13), 3693-
848 3698.
849 Gaspar, J. M., & McDonald, J. J. (2014). Suppression of salient objects prevents distraction in
850 visual search. *Journal of Neuroscience*, *34*(16), 5658-5666.
851 Gaspelin, N., & Luck, S. J. (2018a). Combined electrophysiological and behavioral evidence for
852 the suppression of salient distractors. *Journal of Cognitive Neuroscience*, *30*(9), 1265-
853 1280.
854 Gaspelin, N., & Luck, S. J. (2018b). Distinguishing among potential mechanisms of singleton
855 suppression. *Journal of Experimental Psychology: Human Perception and Performance*,
856 *44*(4), 626-644.
857 Green, J. J., Conder, J. A., & McDonald, J. J. (2008). Lateralized frontal activity elicited by
858 attention-directing visual and auditory cues. *Psychophysiology*, *45*(4), 579-587.
859 Hickey, C., Di Lollo, V., & McDonald, J. J. (2009). Electrophysiological indices of target and
860 distractor processing in visual search. *Journal of Cognitive Neuroscience*, *21*(4), 760-
861 775.

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience*, *18*(4), 604-613.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194-203.

Jannati, A., Gaspar, J. M., & McDonald, J. J. (2013). Tracking target and distractor processing in fixed-feature visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1713-1730.

Kerzel, D., & Burra, N. (2020). Capture by context elements, not attentional suppression of distractors, explains the P_D with small search displays. *Journal of Cognitive Neuroscience*, *32*(6), 1170-1183.

Luck, S. J. (2012). Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 329-360). Oxford University Press.

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition*, *29*, 1-21.

Luck, S. J., Girelli, M., McDermott, M. T., & Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology*, *33*(1), 64-87.

Luck, S. J., & Hillyard, S. A. (1994a). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291-308.

Luck, S. J., & Hillyard, S. A. (1994b). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000-1014.

Maljkovic, V., & Nakyama, K. (1996). Priming of pop-out: II. The role of position. *Perception & Psychophysics*, *58*(7), 977-991.

888 Mazza, V., & Caramazza, A. (2011). Temporal brain dynamics of multiple object processing:
889 The flexibility of individuation. *PloS one*, *6*(2), e17453.
890 Miller, J., Patterson, T. U. I., & Ulrich, R. (1998). Jackknife-based method for measuring LRP
891 onset latency differences. *Psychophysiology*, *35*(1), 99-115.
892 Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential
893 and current density mapping. *Electroencephalography and Clinical*
894 *Neurophysiology*, *72*(2), 184-187.
895 Phipson, B., & Smyth, G. K. (2010). Permutation P-values should never be zero: Calculating
896 exact P-values when permutations are randomly drawn. *Statistical Applications in*
897 *Genetics and Molecular Biology*, *9*(1).
898 Potts, G. F., Liotti, M., Tucker, D. M., & Posner, M. I. (1996). Frontal and inferior temporal
899 cortical activity in visual target detection: Evidence from high spatially sampled event-
900 related potentials. *Brain Topography*, *9*(1), 3-14.
901 Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain and*
902 *Cognition*, *56*(1), 5–13.
903 Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for
904 accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225-
905 237.
906 Sawaki, R., Geng, J. J., & Luck, S. J. (2012). A common neural mechanism for preventing and
907 terminating the allocation of attention. *Journal of Neuroscience*, *32*(31), 10725-10736.
908 Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons:
909 Electrophysiological evidence for an automatic attend-to-me signal. *Attention,*
910 *Perception, & Psychophysics*, *72*(6), 1455-1470.
911 Smith, J. L., Smith, E. A., Provost, A. L., & Heathcote, A. (2010). Sequence effects support the
912 conflict theory of N2 and P3 in the Go/NoGo task. *International Journal of*
913 *Psychophysiology*, *75*(3), 217-226.

914 Stilwell, B., & Gaspelin, N. (in press). Attentional suppression of highly salient color singletons.
915 *Journal of Experimental Psychology. Human Perception and Performance*.
916 Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive
917 waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and*
918 *Clinical Neurophysiology*, *38*(4), 387-401.
919 Tay, D., Harms, V., Hillyard, S. A., & McDonald, J. J. (2019). Electrophysiological correlates of
920 visual singleton detection. *Psychophysiology*, *56*(8), e13375.
921 Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception &*
922 *Psychophysics*, *50*(2), 184-193.
923 Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception &*
924 *Psychophysics*, *51*(6), 599-606.
925 Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta*
926 *Psychologica*, *135*(2), 77-99.
927 Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). On the time course of top-down and bottom-
928 up control of visual attention. In S. Monsell & J. Driver (Eds.), *Control of Cognitive*
929 *Processes, Attention & Performance, XVIII* (pp. 105-124). MIT Press.
930 Tillman, C. M., & Wiens, S. (2011). Behavioral and ERP indices of response conflict in Stroop
931 and flanker tasks. *Psychophysiology*, *48*(10), 1405-1411.
932 Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive*
933 *Psychology*, *12*(1), 97-136.
934 Treisman, A.M., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In
935 S. Dornic (Ed.), *Attention and Performance VI* (pp. 333-361). Erlbaum.
936 Wang, B., & Theeuwes, J. (2020). Salience determines attentional orienting in visual
937 selection. *Journal of Experimental Psychology: Human Perception and*
938 *Performance*, *46*(10), 1051.

939 Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention
940 and how do they do it?. *Nature Reviews Neuroscience*, *5*(6), 495-501.
941 Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of
942 attention during visual search. *Nature*, *400*(6747), 867-869.
943

944 **Figure Legends**

945 **Figure 1.** Example stimulus displays used in Experiments 1 and 2.

946 **Figure 2.** Behavioural results of Experiment 1. Violin plots showing the median, quartiles, and
947 individual-participant data points for RTs of target-present trials, separately plotted based on
948 whether the previous display (trial N-1) contained a target or distractor singleton in the same or
949 different quadrant. Participants were statistically faster to respond to targets appearing at the
950 same quadrant than to targets appearing elsewhere, but there was no evidence of such
951 location-priming effect when the previous trial contained a distractor.

952 **Figure 3.** Grand-averaged singleton-present ERPs recorded over the lateral occipital scalp
953 (electrodes PO7/8) in Experiment 1. Positive voltages were plotted downward by convention,
954 thereby making the N2pc to appear above baseline and the P_D to appear below baseline. (**a**)
955 Left: waveforms recorded contralateral and ipsilateral to singletons on all singleton-present Go
956 trials. Right: waveforms recorded contralateral and ipsilateral to singletons on all singleton-
957 present No-Go trials. (**b**) Contralateral-minus-ipsilateral difference waves corresponding to the
958 waveforms in the previous panel. Vertical bars correspond to the 95% CIs at each time point.

959 **Figure 4.** Grand-averaged waveforms recorded over the midline frontal and central scalp (FPz
960 and Cz, respectively) in Experiment 1. Positive voltages were plotted downward by convention.
961 (**a**) Waveforms elicited by Go and No-Go trials. (**b**) Go-minus-No-Go difference waves
962 corresponding to the waveforms in the previous panel. Vertical bars correspond to the 95% CIs
963 at each time point.

964 **Figure 5.** Isolation of singleton-related processing in Experiment 1. Positive voltages were
965 plotted downward by convention. (**a**) Grand-averaged waveforms of singleton-present
966 (ipsilateral PO7/8) and singleton-absent trials (combined PO7/8). (**b**) Present-minus-absent

967 difference waves recorded over the ipsilateral and contralateral scalp constructed by subtracting
968 the singleton-absent ERPs from the corresponding singleton-present ERPs in the previous
969 panel. Vertical bars correspond to the 95% CIs at each time point. (**c**) Topographic maps of the
970 P3b elicited by singleton-present and singleton-absent displays on Go trials. (**d**) Topographic
971 maps of the present-minus-absent difference wave elicited on Go trials. Whereas the P3b is
972 maximal over the midline parietal scalp (at electrode Pz), the SDP is maximal over the bilateral
973 occipital scalp (at electrodes PO7/8).

974 **Figure 6.** Grand-averaged singleton-present ERPs recorded over the lateral occipital scalp
975 (electrodes PO7/8) in Experiment 2. Positive voltages were plotted downward by convention. (**a**)
976 Waveforms recorded contralateral and ipsilateral to singletons on all target-present trials. (**b**)
977 Comparison of grand-averaged, contralateral-minus-ipsilateral difference waves elicited by
978 target singletons between Experiment 1 and Experiment 2. (**c**) Grand-averaged, present-minus-
979 absent difference waves recorded over the ipsilateral and contralateral scalp. Vertical bars
980 correspond to the 95% CIs at each time point.

981 **Figure 7**. Results of Experiment 3. Violin plots showing the median, quartiles, and individual-
982 participant data points for RTs of 8- and 16-item search displays (Go trials).

983

Figure 1.

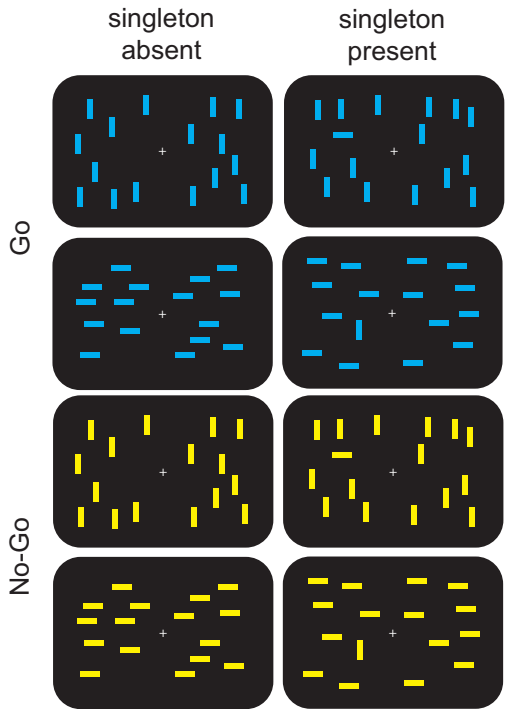
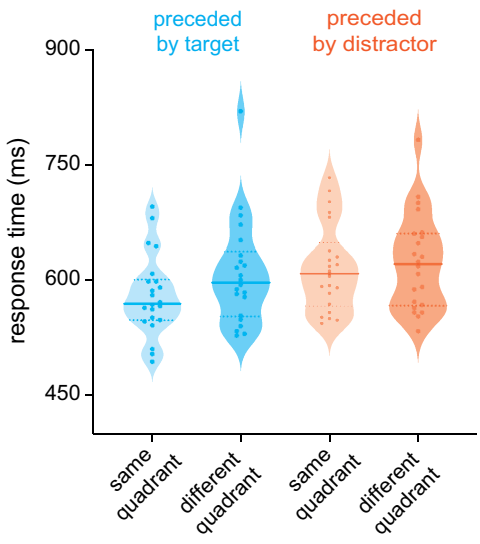
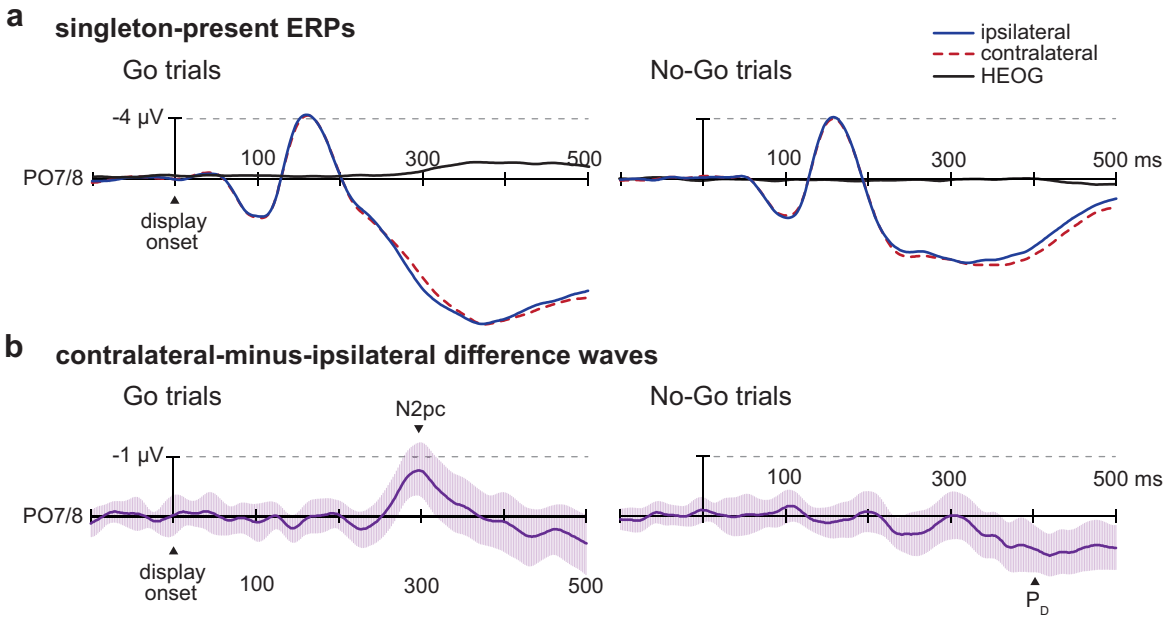


Figure 2.

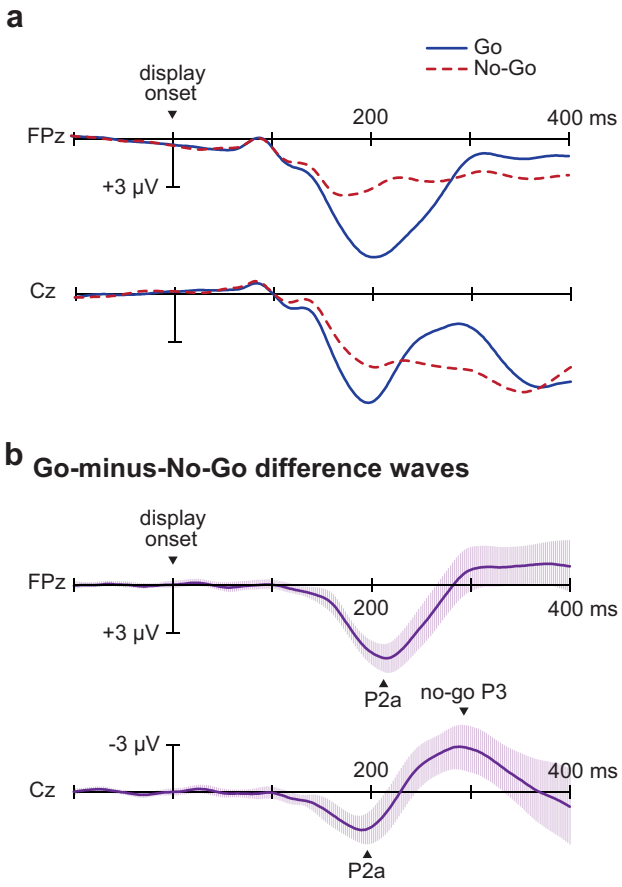

preceded by target
preceded by distractor
response time (ms)
900
750
600
450
same quadrant
different quadrant
same quadrant
different quadrant

988 **Figure 3.**



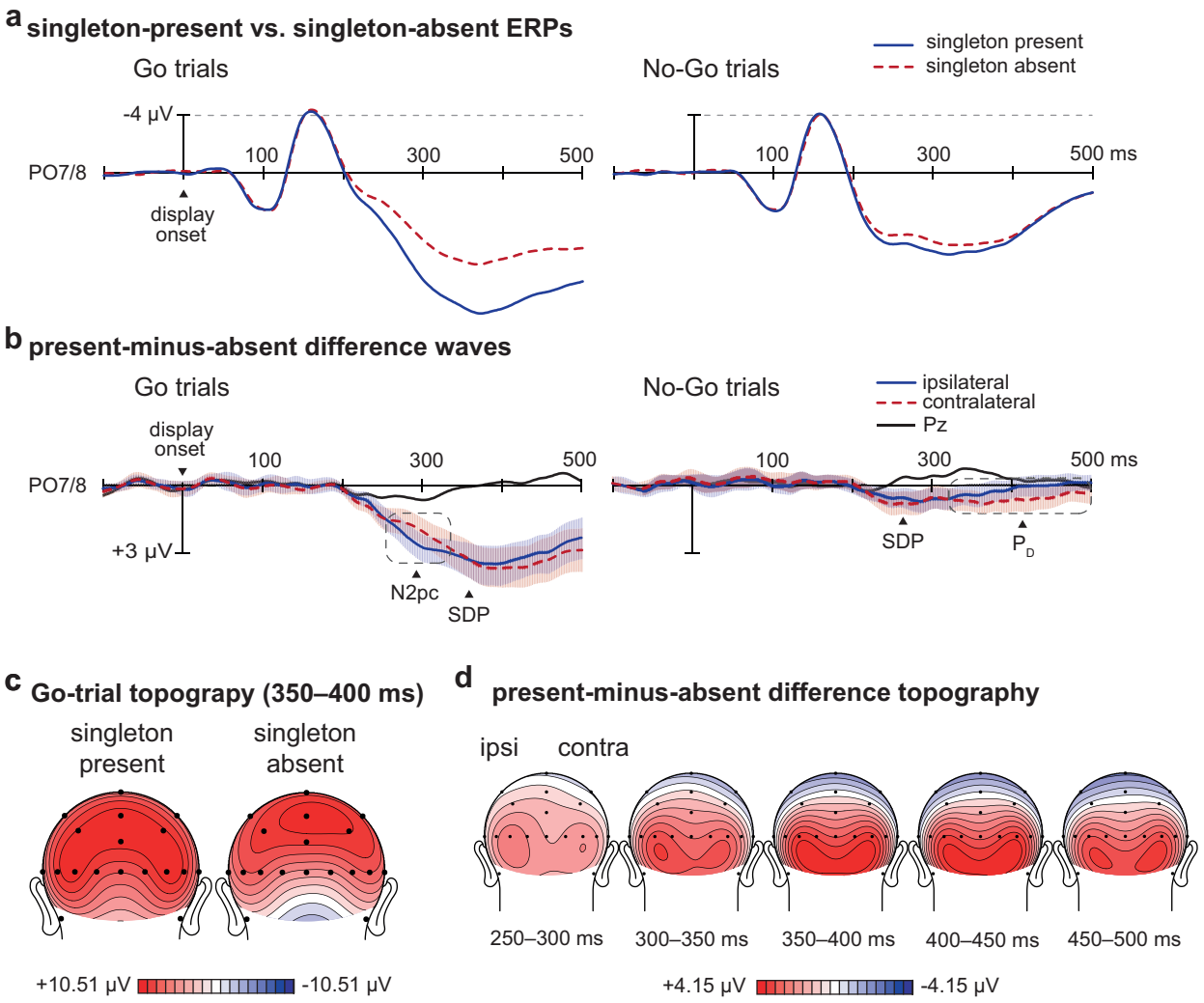
989

990 **Figure 4.**



991

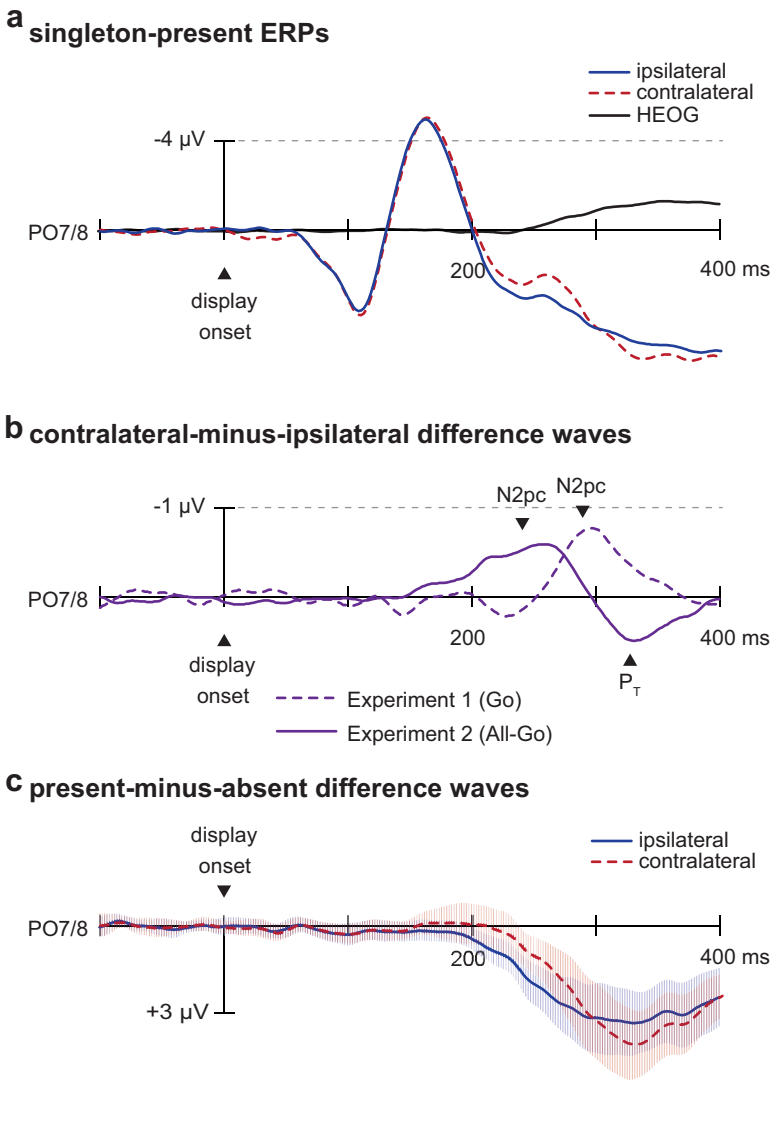
992 **Figure 5.**



993

994

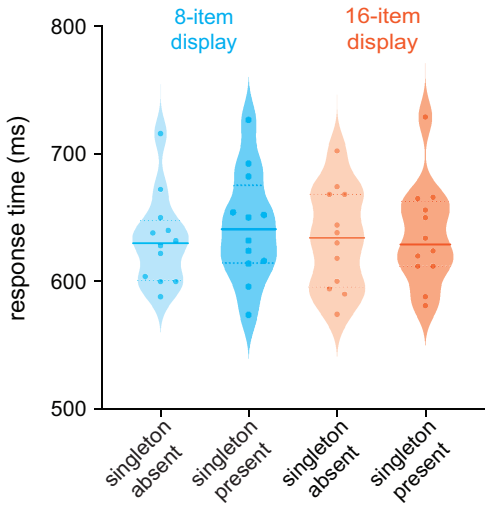
995 **Figure 6.**



996

997

998 **Figure 7.**



999