Journal of Experimental Psychology: General

A Neural Habituation Account of the Negative Compatibility Effect

Len P. L. Jacob, Kevin W. Potter, and David E. Huber

Online First Publication, May 20, 2021. http://dx.doi.org/10.1037/xge0001055

CITATION
Jacob, L. P. L., Potter, K. W., & Huber, D. E. (2021, May 20). A Neural Habituation Account of the Negative Compatibility Effect. *Journal of Experimental Psychology: General*. Advance online publication. http://dx.doi.org/10.1037/xge0001055

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

© 2021 American Psychological Association
ISSN: 0096-3445

https://doi.org/10.1037/xge0001055

A Neural Habituation Account of the Negative Compatibility Effect

Len P. L. Jacob, Kevin W. Potter, and David E. Huber
Department of Psychological and Brain Sciences, University of Massachusetts Amherst

The negative compatibility effect (NCE) is the finding of slower reaction times (RTs) to report the direction of a target arrow following a subliminal prime arrow pointed in the same direction. The NCE is commonly thought to reflect automatic response inhibition, and on this assumption, it has recently been used to assess various motor disorders. Here we propose a fundamentally different account of the NCE: one that relates the NCE to a broader class of paradigms that reveal behavioral deficits with repetition priming. We propose that the NCE is a "cognitive aftereffect," as explained with the neural habituation model of Huber and O'Reilly (2003). To identify the underlying perceptual dynamics by reducing the role of response preparation, we developed a novel variant of the NCE task with threshold accuracy rather than RT as the dependent measure. This revealed a transition from positive to negative priming as a function of prime duration, and a second experiment ruled out response priming. The perceptual dynamics of the neural habituation model were fit to these results and then fixed in applying the model to the NCE literature. Application of the model to RTs added a response layer that accumulates response information throughout the trial. With this addition, the model captured results found in the NCE literature that are inconsistent with a response inhibition account. Situations that produce a positive compatibility effect, rather than an NCE, were explained as response priming, whereas NCE effects were explained as a cognitive aftereffect, rooted in perceptual dynamics.

Keywords: negative compatibility effect, computational model, priming, neural habituation

Muscle contractions produce behavioral responses and neurons trigger muscle contractions. This causal chain suggests that human behavior benefits from being studied in terms of neural behavior, and theories grounded in neural properties have the potential to generalize across different tasks. However, many theories of human behavior are presented at David Marr's (Marr & Poggio, 1977) "computational" level, without reference to neural behavior. Instead, these theories appeal to environmental regularities, task demands, and optimal inference (Griffiths et al., 2008; Huber et al., 2001). Because each task entails different task demands and environmental regularities, this may result in theoretical models with a narrow explanatory scope. More common in cognitive psychology are "algorithmic" theories based on unobservable representations and processes, with many of these theories likewise developed without reference to neural behavior (McKoon & Ratcliff, 1992). Although these computational and algorithmic explanations have been successful in their separate applications to behaviors of interest, they may be overly complex in a larger sense, failing to reveal commonalities between different behavioral tasks—commonalities that reflect a common neural behavior.

By focusing on general properties of neural behavior, we seek to provide unified accounts of different experimental paradigms. To this end, we have identified a form of neural behavior that may be common to a wide variety of experimental effects that involve the rapid presentation of stimuli, examining the short-term effect of a recent stimulus (e.g., a prime or a first target) on a subsequent target stimulus. The neural behavior hypothesized to underlie these effects is synaptic depression (Abbott et al., 1997), which describes a transient reduction in connection strength between a sending neuron and a receiving neuron. Studies indicate that the majority of excitatory pyramidal cells exhibit this behavior (Zucker & Regehr, 2002) and, as such, this may be a core aspect of the link between neural behavior and cognition. There are many possible causes of synaptic depression, although the dominant cause is neurotransmitter depletion due to recent neural activity (Singer & Diamond, 2006). Synaptic depression can produce an order of magnitude reduction in the neural response, and this effect can build up on a time scale of 10s to 100s of milliseconds (Tsodyks & Markram, 1997), in other words, the time scale of perception and action. As a result, when a stimulus is repeated with a short delay between repetitions, the neural response to the second

---

Len P. L. Jacob https://orcid.org/0000-0002-4783-1368
David E. Huber https://orcid.org/0000-0002-7709-7993
Kevin W. Potter is now at the Center for Addiction Medicine, Massachusetts General Hospital, Boston, Massachusetts, United States.

We have no known conflict of interest to disclose.

The ideas and data contained in this work are scheduled to be presented at the 2020 Annual Meeting of the Psychonomic Society.

We provide the deidentified data analyzed in this study along with paradigm and model code in the following OSF repository: https://osf.io/uzy7k/?view_only=e39dbc702482433783647e7e11940dca.

We thank Andrea Mayoral for data collection on a pilot version of the experiment.

Correspondence concerning this article should be addressed to Len P. L. Jacob, Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Tobin Hall, 130 Hicks Way, Amherst, MA 01003, United States. Email: ljacob@psych.umass.edu

occurrence is substantially reduced—so called "stimulus specific adaptation" (Ulanovsky et al., 2004).

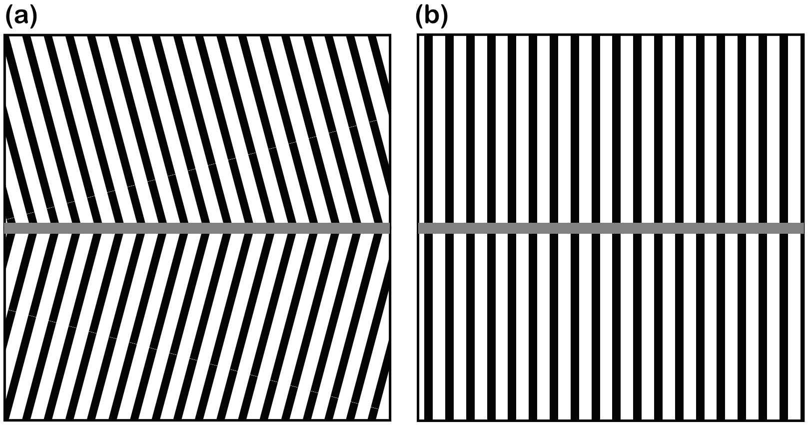
Huber and O'Reilly (2003) proposed that short-term synaptic depression serves an important cognitive function: It parses the rapidly changing stream of visual objects by reducing the neural response to the features contained in previously viewed objects. This allows unobstructed perception of subsequent objects that are comprised of different features, as it reduces the misattribution of features from the previous stimulus to the current stimulus. Huber and O'Reilly termed this transient reduction in the neural response "accommodation," but based on feedback from colleagues and reviewers who found this terminology confusing, all subsequent publications of the theory used the term "habituation." Critically, this habituation is different than longer-term learning and is decidedly different than active inhibition. Instead, it reflects a passive and temporary weakening of the current neural response, as well as a short-term deficit (a few seconds), should a subsequent stimulus involve that same neural response. However, this mechanism for temporal parsing comes at a cost, producing repetition deficits. Huber and O'Reilly derived a rate-coded neural habituation model of synaptic depression from the spiking-neuron model of Chance et al. (1998), and successfully applied this model to behavioral priming effects.

Since development of the neural habituation model in 2003, it has explained a wide variety of behavioral paradigms that examine short-term interstimulus effects. Analogous to a visual aftereffect, such as the tilt aftereffect (see Figure 1), these effects can be thought of as "cognitive aftereffects" (i.e., aftereffects for perceptual identification representations that are at a higher-level as compared with retinotopic visual properties). For example, after viewing a figure that is white, an observer will experience a black afterimage (a visual aftereffect) and, analogously, after reading the letter X, the reading system is better able to "see" the letter O, even if the X and O appear in different visual locations (a cognitive aftereffect). These applications of the neural habituation model include repetition and semantic priming of words (Huber, Tian, et al., 2008; Jacob & Huber, 2020; Potter et al., 2018; Rieth & Huber, 2017), repetition priming of faces (Rieth & Huber, 2010), change detection with words (Davelaar et al., 2011), semantic satiation (Tian & Huber, 2013), inhibition of return with spatial cuing (Rieth & Huber, 2013), priming of affective valence (Irwin et al., 2010), immediate repetition in recognition memory (Huber, Clark, et al., 2008), and temporal search for targets in the "attentional blink" paradigm (Rusconi & Huber, 2018). The common link between the aforementioned paradigms is the presence of a rise-then-fall behavioral effect: brief presentations (or brief interstimulus intervals) yield positive effects, such as faster or more accurate target identification, but as presentations (or interstimulus intervals) become longer, deficits occur instead. This shift is attributed to increased neural habituation, which reduces the neural response of the features and identity of the previously seen stimuli.

Here, we add to the list of paradigms explained by neural habituation and cognitive aftereffects, demonstrating that a long-standing debate surrounding the causes of the "negative compatibility effect," or NCE (Eimer & Schlaghecken, 1998) can be resolved with this model. In doing so, we further unify the literature of short-term interstimulus effects, bringing it together under the shared neural behavior of short-term synaptic depression.

The NCE paradigm presents arrows as a prime stimulus followed by arrows as a target stimulus, finding slower responses (i.e., a negative aftereffect) to the target when the direction of the prime arrow is the same (compatible) with the target. As reviewed below, it has been debated whether the NCE is a perceptual effect or a response effect. However, as response arises from perception, using the labels perception or response is not always clear; furthermore, these are not necessarily mutually exclusive explanations, as it is possible for both perceptual and response priming to generate effects on behavior within the same task, sometimes even in opposite directions. For instance, a recently viewed stimulus can produce both a negative visual aftereffect, biasing perception away from that stimulus, and yet, at the same time, the decision process might be biased to give the same answer that was given to that prior stimulus (Pascucci et al., 2018). The approach we take here

Figure 1
Tilt Aftereffect



Note. Fixating on the horizontal gray line in (a) for 10 s, then switching fixation to the horizontal gray line in (b), causes the vertical lines in (b) to be perceived as slightly tilted in the opposite direction as the lines in (a; Gibson & Radner, 1937).

is to isolate the purely perceptual components of the NCE by using a modified version of the task with threshold accuracy as the dependent measure, rather than speeded reaction times. Then, with this perceptual component fixed, we apply our theory to the typical NCE task to specify the additional role of decisional processes. To preview our results, although we conclude that that response priming plays a role in the typical reaction time (RT) NCE task, the nature of the response effect is positive rather than negative, whereas the often observed negative effects in the NCE task reflects a higher-level perceptual aftereffect (i.e., a cognitive aftereffect for the general sense of arrow direction).

The Direction of Cognitive Aftereffects as Explained by Neural Habituation

There are many levels of processing in the ventral visual stream for identifying objects, ranging from line segments to complex conjunctions to possible responses (Kobatake & Tanaka, 1994), and the negative effects of a recent stimulus may occur at any or all of these levels. The tilt aftereffect (Gibson & Radner, 1937; see Figure 1) is an example of a negative aftereffect at a low-level. After fixating on the horizontal gray line in the left panel of Figure 1 (Figure 1a) for several seconds, you may perceive the lines in the right panel (Figure 1b) as slightly tilted in the opposite direction (in laboratory experiments the effect is stronger because the second display abruptly replaces the first, rather than requiring a saccade between the two displays). While this particular effect is primarily retinotopic (Jin et al., 2005), similar high-level negative aftereffects can occur irrespective of retinotopic position (Fox, 1995). Such aftereffects can affect cognition and decision-making, priming subjects against the properties of recently viewed stimuli. However, the opposite can also happen, with a recent stimulus causing a bias to respond *with* the previously seen stimulus (Fischer & Whitney, 2014). What then determines whether a previously seen stimulus generates a positive or negative bias?

In a word priming paradigm, Huber, Tian, et al. (2008) found that increasing prime duration from tens of milliseconds to seconds causes a transition from a positive effect to a negative effect for the case of repetition priming between a prime word and a target word presented immediately after the prime. The neural habituation model explained these results as emerging from a multilevel hierarchical system for word identification, with neural habituation dynamics occurring at all levels. When considering a level in isolation, the response to a stimulus at that level first increases, due to temporal integration, and then decreases, due to neural habituation.

In repetition priming with words, the letters and meaning of the prime reappear in the target (prime and target were visually distinct by being presented in different locations). In other words, in the multilevel hierarchical system, different simulated neurons in the retinotopic layers are activated for the prime and the target, but if these words are the same, both neuronal populations activate the same letter/meaning representations. Following a briefly presented prime (e.g., 50 ms or less), lingering activation for the orthography of the prime boosts the neural response to the target word (i.e., it is easier to identify the letters of the target word because the prime gives those letter identities a Head Start). This produces positive priming. In contrast, following a long duration prime (e.g., 500 ms or more), synaptic depression reduces the magnitude of this neural response by weakening the connections between the neural populations that represent orthography and those that represent meaning. Negative priming emerges as a relative deficit in the neural habituation model, similar to the Bayesian discounting found in the responding optimally with unknown sources of evidence (ROUSE) model (Huber et al., 2001). Because synaptic resources take time to recover, these connections are still weakened at the time when the target is presented. Thus, in response to a briefly presented target word that repeats the long duration prime, even though there is a small benefit of lingering activation from the prime, this is more than offset by habituation, and the word identification system struggles to extract the meaning of the target word from its letters, producing negative priming (worse performance as compared to what would have occurred in the absence of a prime).

This shift from positive to negative effects as described above is observed in all of the paradigms explained by the neural habituation model (as outlined in the previous section). When modeling each task, a similar multilevel hierarchical system is used, beginning with a retinotopic layer that feeds into increasingly more complex representations (such as facial features, then whole faces). Thus, by adjusting the layer architecture and the node structure of the retinotopic layer in a manner that is appropriate to the stimuli of each task, the model can be extrapolated to explain a variety of tasks and stimuli. The theoretical link between these tasks is not in terms of perceptual representations, but rather shared neural dynamics.

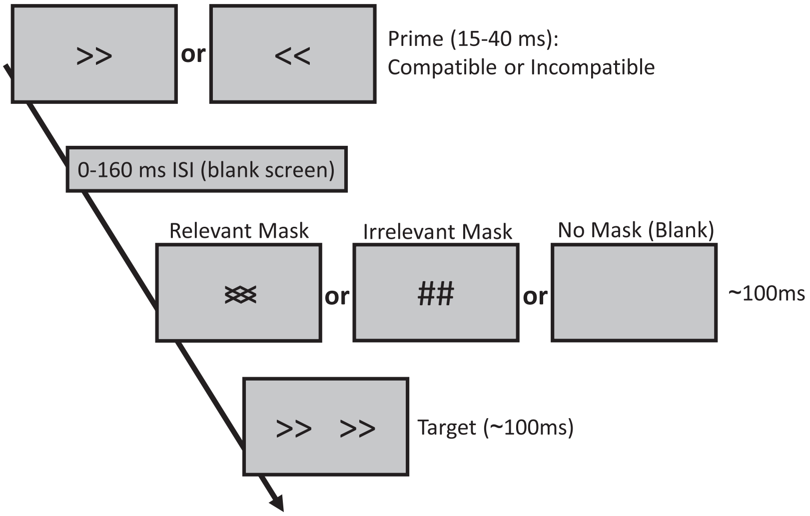
The Negative Compatibility Effect

Similar to repetition priming with words, the negative compatibility effect involves presentation of a first stimulus that produces a behavioral deficit for a subsequent target stimulus when the key property of the target (i.e., whether the target is oriented to the right or left) repeats the first stimulus. This was first described by Eimer and Schlaghecken (1998), who had subjects perform a simple arrow priming task (see Figure 2). Each trial began with a prime consisting of two left- or right-pointing arrows, immediately followed by a mask composed of two left-pointing arrows superimposed on two right-pointing arrows (the "Relevant Mask" in Figure 2). Finally, a target display was presented, and subjects were instructed to identify the direction of the arrows in the target display by pressing corresponding buttons (one to their right, pressed with their right hand; one to their left, pressed with their left hand). Surprisingly, if the prime direction was compatible with the target direction, reaction times to the target were significantly slowed. This was termed the negative compatibility effect. Because the NCE was first described, several researchers have replicated the original effect (Jaskowski & Przekoracka-Krawczyk, 2005; Klapp & Hinkley, 2002; Lleras & Enns, 2004), in addition to expanding upon the original paradigm to study the effect of different mask types.

Figure 2 shows a typical arrow priming paradigm used to study the NCE. The prime is presented at the same screen position as the mask (when a mask is present), while the target is doubled-up and slightly offset so that it does not share a screen position with the masked prime. Prime duration and prime-mask interstimulus interval (ISI) vary across studies but are typically in the 15–40 ms and 0–160 ms ranges, respectively. Mask and target are displayed for roughly 100 ms, with minor variations. Masks can be "relevant"

Figure 2

Representative Arrow Priming Paradigm Used in NCE Studies



Note. Subjects must identify target direction by pressing corresponding buttons with their right or their left hand. Timing and mask types varied across studies. Eimer and Schlaghecken (1998) used 17 ms primes, 0 ms interstimulus interval (ISI), 117 ms relevant masks, and 133 ms targets. Relevant masks include line segments in the same angles as the primes, while irrelevant masks consist of vertical and horizontal lines. NCE = negative compatibility effect.

(include line segments in the same angles as the primes) or "irrelevant" (do not include line segments in the same angles as the prime; typically consist of vertical and horizontal lines).

Much of the later work in the NCE literature focused on changes in the NCE when comparing different mask types, leading to opposing theories of the NCE, differing primarily on the basic question of whether the NCE reflects perceptual or response processes. The two main accounts of the NCE are known as the self-inhibition account (Schlaghecken et al., 2008), which appeals to responses processes, and the object updating account (Lleras & Enns, 2004), which appeals to perceptual processes.

According to the self-inhibition account, the mask triggers self-inhibitory circuits that cause the ongoing response to the prime to be suppressed (Eimer & Schlaghecken, 1998; Klapp & Hinkley, 2002; Schlaghecken et al., 2008). An EEG signal related to response processes (the lateralized readiness potential or LRP) seemingly supported the self-inhibition account (Eimer & Schlaghecken, 1998; Liu et al., 2014; Praamstra & Seiss, 2005). More specifically, LRPs to the prime indicated activation of the response that matched the prime (activation of the correct response for compatible primes, and of the incorrect response for incompatible primes), but the LRPs reversed in response to the mask, which was attributed to self-inhibition. A key aspect according to the self-inhibition account is that the prime is presented below the threshold of conscious perception, while still generating enough activation to cross a theoretical threshold that can trigger self-inhibitory circuits (Eimer & Schlaghecken, 2002, 2003; Klapp & Hinkley, 2002; Schlaghecken & Eimer, 2002). Thus, according to this account, the NCE is reliant on the mask being present and capable of suppressing prime identification, regardless of mask composition (Eimer & Schlaghecken, 2002; Klapp, 2005; Klapp & Hinkley, 2002; Schlaghecken et al., 2008). Therefore, the prime must be masked for the NCE to occur, and provided that the mask successfully blocks awareness of the prime's identity, the NCE should occur regardless of mask type (relevant or irrelevant). If the mask cannot suppress prime identification—thus, if the prime is supraliminal instead of subliminal—the account predicts a positive compatibility effect (PCE—faster reaction times when the prime matches the target) instead of NCE (Bowman et al., 2006).

The object updating account was first proposed by Lleras and Enns (2004). According to this account, perception of the mask interacts with the prime, and if they share features (i.e., if the mask is relevant), the new features of the mask "pop out" (e.g., after viewing rightward slanted lines in the prime, followed by a mask with both right-ward and left-ward slanted lines, the left-ward lines of the mask are perceptually salient). Therefore, if the prime is compatible with the target, the salient features of the mask will be incompatible with the target, slowing identification of the target and causing the NCE. According to this account, the NCE only occurs with relevant masks, but not with irrelevant masks or in the absence of a mask (Lleras & Enns, 2004, 2005).

The literature has supported both accounts to some degree, but, at the same time, key predictions of both accounts have been falsified. The self-inhibition account's requirement of subliminal primes has been disproven by Lleras and Enns (2004), who found that the NCE still occurred when subjects could identify the primes, and studies have found that the use of irrelevant masks can cause a PCE instead of NCE (Lleras & Enns, 2004, 2005). On the other hand, the object updating account's prediction that irrelevant masks should fail to produce an NCE has been disproven (Jaskowski, 2008; Klapp, 2005; Schlaghecken et al., 2008). Finally, neither account can explain the presence of the NCE when the mask is replaced by flankers

(Jaskowski, 2008) or when there is nothing but a blank screen between prime and target (Klauer & Dittrich, 2010).

Despite the lack of consensus regarding the cause of the NCE, recent studies of the NCE have reverted to the original self-inhibition explanation; several studies have assumed a response inhibition locus and employed NCE paradigms to study neurological motor disorders (D'Ostilio et al., 2013; Rawji et al., 2020; Stenner et al., 2018; Sumner et al., 2007). In these studies, the absence of an NCE is taken as evidence of damaged motoric response inhibition circuits in patients with lesions (Sumner et al., 2007), Parkinson's disease (D'Ostilio et al., 2013), or tic disorders (Rawji et al., 2020; Stenner et al., 2018). This adoption of the original response explanation likely stems from the event-related potential (ERP) results, which clearly show response preparation in the direction opposite to the prime. However, as we describe in the General Discussion section, the finding that response signals track the NCE does not necessarily indicate that they are the cause of the NCE—the motor system can only respond to things that have first been perceived, and it might be that the response signals are merely tracking perceptual signals. Thus, it may be inaccurate and possibly misleading to use the presence or absence of the NCE as an indicator of whether motor response circuits are adequately functioning. Perhaps the authors of these studies (Rawji et al., 2020; Stenner et al., 2018), did not consider alternative explanations of the NCE because the alternatives have failed to provide a comprehensive account of the variable nature of the NCE with different manipulations. Here we resolve this issue, demonstrating that an alternative account of the NCE, with its roots in perception rather than response preparation, can provide a comprehensive explanation of the situations that do, or do not, produce an NCE.

Beyond self-inhibition and object updating, alternative explanations for the NCE have been suggested (Jaskowski, 2008; Klauer & Dittrich, 2010), but these accounts do not provide a mathematical/computational model that handles all of the major findings with the same set of parameter values. The major findings to explain include: a universal finding of NCE with relevant masks (regardless of ISI), and either PCE or NCE for both irrelevant masks or no mask, depending on prime-mask ISI (more specifically, a transition from PCE to NCE with increasing prime-mask ISI).

Overview of the Current Study

The finding in the NCE literature that arrow primes followed by an irrelevant mask or no mask produce a transition from PCE to NCE as prime-mask ISI increases is remarkably similar to the shift from positive to negative priming found in prior studies explained by the neural habituation theory. If the dynamics of neural habituation can explain these effects, this will provide a parsimonious account of the NCE, as the neural habituation model also explains a wide variety of other paradigms that involve the rapid presentation of stimuli. Thus, rather than providing an explanation that is unique to the NCE, the NCE literature would be reinterpreted as another example of what happens when the brain attempts to parse a rapid sequence of events by habituating to previously viewed stimuli.

At its core, the neural habituation is a model of perception, with habituation serving to minimize perceptual interference from recently viewed stimuli. However, by using easily seen target stimuli and by using RT as the key dependent measure, the NCE paradigm lends itself to explanations in terms of both perception (e.g., orientation priming) and response (e.g., motor response priming). Because subjects are attempting to respond quickly, they may mistakenly begin to respond to the prime, before inhibiting that response upon realizing their error. To minimize the role of response preparation and inhibition, we used a modified accuracy version of the NCE paradigm in which observers could take their time in responding, with performance limited by presenting impoverished perceptual information. In our variant of the NCE, subjects were instructed to identify the dominant orientation for a target display that contained a pair of orthogonal overlapping gratings (i.e., a "plaid"), and we set visual contrast of the target separately for each subject to produce accuracy close to 75% (i.e., threshold accuracy as the dependent measure, rather than RT).

Our accuracy version of the NCE allowed us to specify the perceptual dynamics of orientation priming. This was achieved with manipulations of prime duration, without any intervening stimuli between prime and target. Thus, we varied prime duration and the prime-target relationship (primes could be compatible, incompatible, or neutral). If the compatibility effect can be explained by neural habituation, it should be heavily dependent on prime duration, and the accuracy results should show a shift from positive priming to negative priming for compatible trials, and the opposite for incompatible trials. We ran two experiments confirming these predictions, with the second ruling out an alternative explanation in terms of response priming by using same/different testing rather than left/right responses.

The neural habituation model was then fit to our accuracy results, specifying appropriate processing-speed parameter values for the perception of orientation. These lower levels of the model were then fixed, and a response layer was added in applying the model to RT results from NCE experiments that used brief primes followed by a mask between prime and target. According to the habituation model, even a brief prime causes some habituation, and this habituation can result in a negative effect when there is a sufficiently long ISI between prime and target. During the ISI, activation from the prime fades (i.e., the cause of positive perceptual priming is lost), but habituation remains (e.g., neurotransmitter may still be depleted for several seconds; the cause of negative perceptual priming remains). The key question asked with our simulation study of the NCE literature was whether the orientation identification parameters derived from our orientation priming experiment could naturally explain ISI effects in the NCE literature, including interactions between ISI and the type of mask presented between prime and target.

Experiment 1a: Orientation Identification

Introduction

We developed an orientation identification task with accuracy as the dependent measure to map out the dynamic time course of orientation perception. At first glance, our task may appear similar to the tilt aftereffect. However, there are some key differences: The tilt aftereffect is retinotopic and thus location-specific; our orientation identification task, on the other hand, presents the prime and target at different positions and of different spatial frequencies such that any effect of the prime on the target likely exists at a higher level than primary visual cortex. Thus, we sought to identify the time course of identification for a more general concept of orientation (alternatively, this can be thought of as orientation perception with large receptive fields). In addition, we included a neutral prime condition to provide a baseline with the same metacontrast masking aspects as occur for the priming conditions (Francis, 1997), without indicating either a compatible or incompatible orientation.

To minimize the role of response preparation, the dependent measure in the orientation identification task was accuracy, in contrast to the RT tasks in the NCE literature. Our modified design should allow us to determine the perceptual dynamics involved in the NCE, which will serve as a basis for Experiment 2, in which decisional components are added upon this perceptual basis for a full account of the effects. Additionally, by using accuracy as the dependent measure, we are able to run parametric manipulations with a larger number of conditions; if we were to investigate RT instead, a greater number of trials per condition would be needed, making it unfeasible to map out the full profile of the timing effects.

To measure accuracy while avoiding ceiling or floor effects, identification perception was placed at threshold separately for each subject based on the visual contrast ratio between the overlapping gratings presented in the target display (one grating at 45°, and the other at 135°, with the higher contrast grating being the correct answer).

Method

Participants

A power analysis consisting of simulations using the R package Superpower (Lakens & Caldwell, 2019) suggested that a sample of 40 subjects would yield power of over 95% for the two-way interaction of interest (prime type and prime duration—see sections below). The parameters used in the simulations were the following: (a) no correlation between measures (chosen as a conservative estimate); (b) predicted standard deviation of .1 (based on a binomial distribution of mean .75—the accuracy value subjects were thresholded to—and 36 samples, the number of trials per condition); (c) predicted means of .75 for all durations when using neutral primes; (d) predicted means of .77, .7525, .735, .7175; and .7 for compatible primes (from shortest to longest duration); and (e) predicted means of .68, .6975, .715, .7325, .75 for incompatible primes (from shortest to longest duration); the mean values were based on previous findings of short-term accuracy priming effects (Huber, Tian, et al., 2008; Huber & O'Reilly, 2003).

To account for potential issues such as equipment malfunction, data from 50 subjects aged 18–35 were collected. Every participant provided written informed consent, and all study procedures were approved by the University of Massachusetts Amherst Institutional Review Board. Volunteers received Psychology course credit as compensation for participating. Subjects reported normal or corrected-to-normal vision. Out of the 50 subjects who participated, three were excluded from all analyses due to accuracy below 60% across all conditions (global accuracy for the three excluded subjects was 49.8%, 56.3%, and 57.4%), 47 participants were included in the analyses.

Materials and Display Sequence

The experimental task (see Figure 3) was displayed on a 24" LCD monitor with a 120 Hz refresh rate and 1080 pixels resolution. Visual stimuli were generated using PsychToolbox (Kleiner et al., 2007) implemented in MATLAB 2015a (The MathWorks Inc, 2015). Each trial began with a fixation cross displayed for 200 ms, followed by a placeholder stimulus outlining where the prime would be displayed. The placeholder stimulus size was 200 pixels (outer circle diameter) and its visual angle was 5°; subsequent stimuli were displayed within these bounds. The placeholder was presented for 400 ms minus prime duration for that particular trial, ensuring a constant duration between onset of the placeholder and onset of the target. This was critical for allowing subjects to know when the target would appear. Following the placeholder, the prime was presented for either 8 ms, 17 ms, 34 ms, 68 ms, or 138 ms. The prime consisted of lines angled at 45°, 135°, or 90°. The 90° prime was neutral (i.e., this orientation never appeared in the target display), the other two angles could be categorized as either compatible or incompatible, depending on target orientation.

The target display was presented immediately after the prime and shown for 84 ms. This display consisted of two parts: a mask in the outer circle (mask of the prime) and the target plaid in the inner circle. The target plaid consisted of two superimposed sets of line gratings: one set at an angle of 45°, the other at an angle of 135°. One grating was of higher contrast (i.e., greater difference between the darker lines and lighter lines): This was the target that subjects were instructed to identify.

Discriminability between the higher contrast grating (from now on referred to as the target) and the lower contrast grating (from now on referred to as the foil) was adjusted by varying the contrast of the target (C_T); foil contrast (C_F) remained constant, fixed at .05. Target contrast was obtained by multiplying the foil contrast by a contrast ratio, which was modulated throughout the experiment to keep subjects at a threshold corresponding to 75% accuracy averaged across all conditions. All contrasts were calculated using Michelson contrast (Pelli & Bex, 2013).

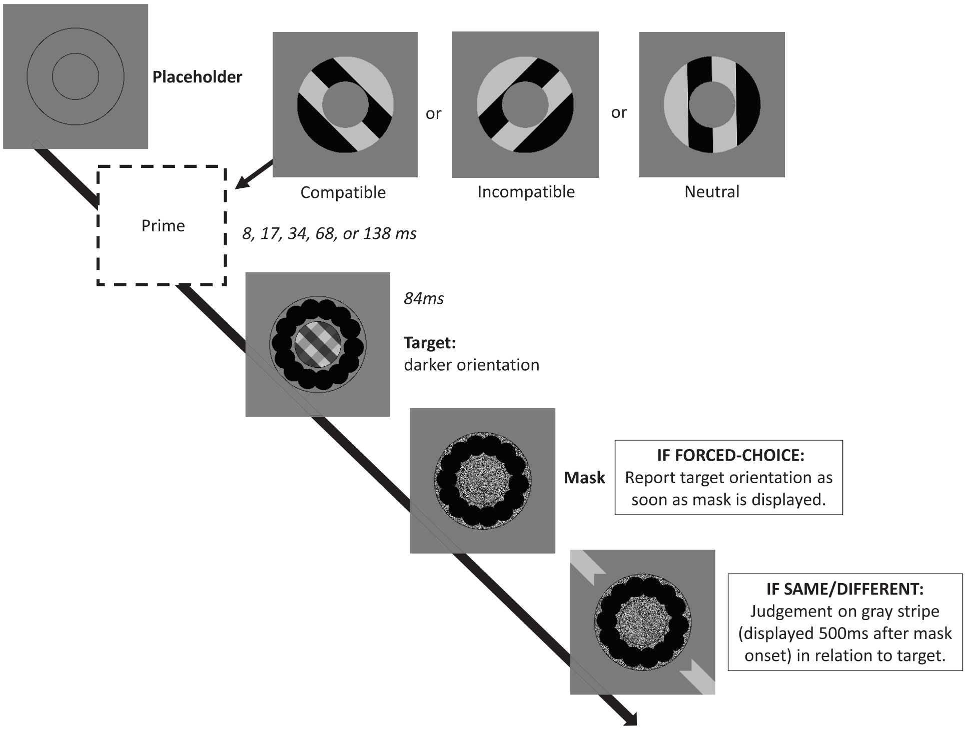
The target display grating stimulus was generated by converting the target and foil contrasts (C_T and C_F) to luminance values (equations shown in Figure 4). To obtain the plaid pattern, four values were needed, in order of brightest to darkest: W (fixed at 76 for the entire experiment), X (calculated using contrast from W and C_F), Y (calculated using contrast from W and C_T), and Z (calculated using W, C_F, and C_T, or simplified to $X*Y/W$). These calculations ensure that the contrasts of both the target and foil are constant regardless of whether the lines are placed against the light or the dark phases of the other orientation. These four values were then gamma corrected to convert luminance values to grayscale values appropriate to the LCD screen (Pelli & Zhang, 1991).

Procedure

Pilot experiments determined that the easiest way to describe the task to subjects was to tell them to determine which direction had darker lines (but we note that this direction also has brighter

Figure 3

Display Sequence for Experiments 1a and 1b



Note. For Experiment 1a (forced-choice), subjects responded by indicating the orientation of the darker set of lines in the target display (right-leaning or left-leaning). For Experiment 1b (same/different), subjects responded by indicating if the outer gray strip was in the same direction as the darker set of lines in the target display. Trials in both experiments began with a fixation cross (not pictured) displayed for 200 ms, followed by the placeholder stimulus displayed for 400 ms minus prime duration. The prime was then displayed for either 8 ms, 17 ms, 34 ms, 68 ms, or 138 ms, followed by a target plaid created from overlapping orthogonal gratings, presented for 84 ms. (For illustrative purposes, in this figure target contrast is higher than it was in the actual experiment.) After target display presentation, a mask was displayed. At this point in Experiment 1a, subjects reported target orientation. In Experiment 1b, the mask was displayed by itself for 500 ms, followed by a gray outer stripe for a same/different response (in this example, the correct response would be "same," as the light gray stripe is of the same orientation as the darker set of stripes within the previously shown target grating).

lines, depending on whether one pays attention to dark vs. light phases of the grating). As soon as the target appeared, subjects could report target orientation by pressing either the F key (for left-leaning targets) or the J key (for right-leaning targets).

The experiment began with 24 practice trials, during which prime duration was always 34 ms, and other properties were counterbalanced (eight compatible trials, eight incompatible trials, and eight neutral trials; target orientation was also counterbalanced with 12 left-learning and 12 right-leaning). The contrast ratio was five for the first 12 trials, and four for the subsequent 12 trials. Practice trials were followed by 120 adjustment trials (with prime duration also fixed at 34 ms, and trials counterbalanced as described above) in which contrast ratio started at 2.75 and was modulated by progressively smaller steps every 24 trials (i.e., a staircase procedure): If subjects were correct on fewer than 16 trials, the contrast ratio was increased, and if they were correct on more than 20 trials, the contrast ratio was reduced. The average contrast ratio of all subjects across all experimental trials was 2.3 (1.1 standard deviation). Step size was .5 for the first accuracy check, then .25 for the next three accuracy checks, then .125 for the last check and the remaining checks that happened during subsequent experimental trials. These 144 practice and threshold trials were not included in the analyses.

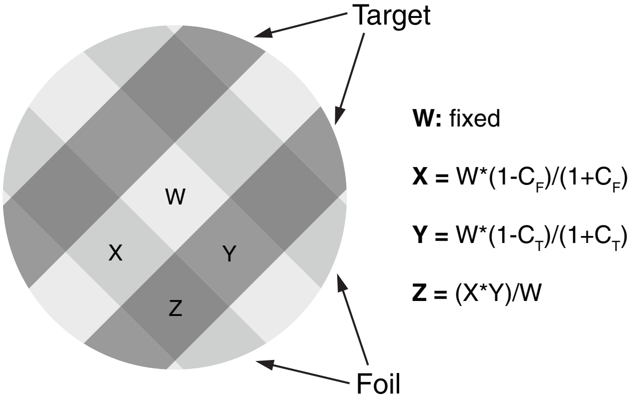
Subjects were provided feedback ("CORRECT" or "WRONG") after every trial for the entire experiment, and trials were self-paced, with the next trial starting only after the subject pressed a key after receiving feedback. This trial-by-trial accuracy feedback was provided to place an emphasis on accurate responding rather than speeded responding. Subjects performed 540 experimental trials (36 trials per prime type/prime duration combination) divided into nine blocks of 60 trials; trials were randomized within blocks so that each block contained four trials per prime type/duration combination. Subjects were notified when a block ended and informed of how many blocks remained.

Results

Analysis of variance was conducted using R and RStudio (RStudio Team, 2020), with prime type and prime duration as factors

Figure 4

Target Display Plaid Created From Four Luminance Values (W, X, Y, and Z), Before Gamma Correction to Determine LCD Grayscale Values



Note. The three equations calculate luminance based on target contrast (C_T), foil contrast (C_T), and a fixed luminance (W) for the brightest squares. These equations ensure that the visual contrast of both the target and foil gratings are constant regardless of whether the grating is placed against the light or dark phases of the other orientation (see text for details).

(with subject number as the error factor). Prime type contained three levels (compatible, incompatible, and neutral) and prime duration contained five levels (8 ms, 17 ms, 34 ms, 68 ms, and 138 ms).

Accuracy across conditions is displayed in Figure 5 Statistical analysis revealed significant main effects of prime duration, $F(4, 184) = 9.57$, $p < .001$, $\eta_p^2 = .172$, and prime type, $F(2, 92) = 12.46$, $p < .001$, $\eta_p^2 = .213$, along with a significant interaction between the two, $F(8, 368) = 26.31$, $p < .001$, $\eta_p^2 = .364$. To identify the crossover point of this significant interaction, we ran uncorrected posthoc pairwise t-tests on the difference between the compatible and incompatible conditions at each prime duration (see Table 1). The difference between means flipped from positive to negative when transitioning from 34 ms to 68 ms primes.

As seen in Figure 5, accuracy in the neutral prime condition is relatively flat across prime durations and generally both the compatible and incompatible conditions were lower or equal to the neutral baseline, with the exception of an 8-ms compatible prime, which produced accuracy that was higher than the neutral prime condition. While perhaps surprising, the finding of deficits for both target repetition priming and foil repetition priming as compared with a baseline condition with unrelated stimuli has been found in many previous word priming experiments (Huber, Shiffrin, Lyle, et al., 2002; Huber, Shiffrin, Quach, et al., 2002; Huber, Tian, et al., 2008; Rieth & Huber, 2017; Weidemann et al., 2005, 2008) and this aspect of the data is predicted by both the Bayesian ROUSE model (Huber et al., 2001) and the neural habituation model (Huber & O'Reilly, 2003). In light of this, we focus on the difference between the compatible and incompatible conditions to assess changes in the direction of priming. As seen in the figure, positive priming effect progressively reversed as prime duration increased, with strong negative priming present at the longest prime duration of 138 ms. Incompatible primes, on the other hand, generated strong negative priming at the shortest prime duration; this negative priming effect progressively weakened as prime duration increased, until it disappeared at longer prime durations as accuracy in the incompatible condition reached similar values as the neutral baseline condition.

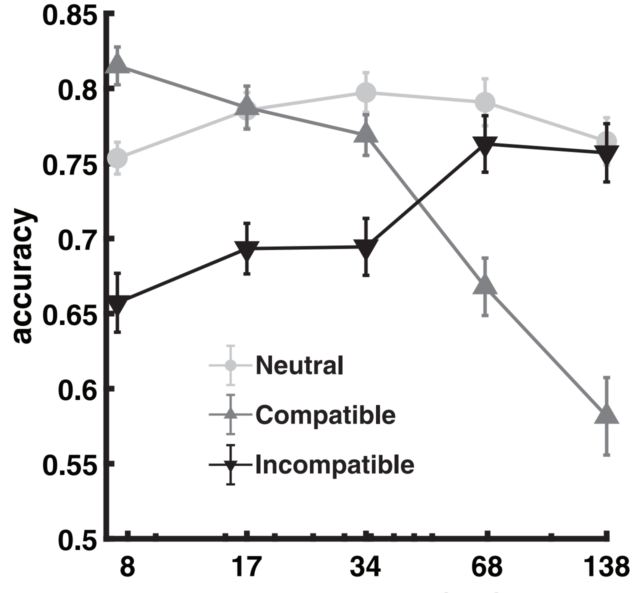
Discussion

The accuracy results are consistent with effects observed in other paradigms explained by the neural habituation theory, showing a shift from positive priming (compatible higher than incompatible) to negative priming (compatible lower than incompatible) as primes are presented for longer. However, the time course of this transition is substantially faster than repetition priming with words, which exhibit a priming crossover between prime durations of 150 ms to 400 ms (Rieth & Huber, 2017), in contrast to the crossover between 34 ms and 68 ms seen here with orientation priming. The fast time scale can be explained by the fact that orientation perception occurs very early in the visual system, while complex stimuli such as words require much longer to be processed. In fact, it has been shown that the tilt aftereffect can be observed with prime displays as short as 18 ms (Sekuler & Littlejohn, 1974).

In terms of the NCE literature, it is important to note that the negative priming effect found here occurred even though there was no mask intervening between prime and target (i.e., an NCE does not require a relevant mask). There was a mask presented

Figure 5

Average Accuracy for Experiment 1a as a Function of Prime Duration (Log Scale) and Priming Condition



Note. Neutral primes consisted of vertical lines displayed at 90°. Compatible refers to primes that matched the target angle (45° or 135°), and incompatible refers to primes that did not match the target angle and instead matched the foil angle (135° or 45°). Error bars are plus and minus one standard error of the mean (2 *SEM*s in total).

Table 1

Paired t-Tests Comparing Accuracy for the Compatible and Incompatible Conditions for Experiment 1a (df = 46)

Prime duration	8 ms	17 ms	34 ms	68 ms	138 ms
t-statistic	6.3	4.67	3.71	−3.24	−4.66
p-value	$p < .0001$***	$p < .0001$***	$p = .00056$***	$p = .0022$**	$p < .0001$***
Mean diff.	0.158	0.094	0.074	−0.095	−0.176

** $p < .01$. *** $p < .001$.

after the prime, but that mask was concurrent with the target and was task irrelevant (a pattern of circles). Much of the NCE literature has focused on the effects of different intervening mask types as a method for differentiating between competing theories (Jaskowski & Przekoracka-Krawczyk, 2005; Klapp, 2005; Lleras & Enns, 2004, 2005, 2006; Schlaghecken et al., 2008); however, as the present results show, whether PCE or NCE occurs can be purely determined by prime duration, with no intervening mask.

While the NCE literature has used speeded responses and RTs as a dependent measure, this orientation paradigm removed any pressure to respond quickly by placing perceptual identification at the accuracy threshold of 75% and by providing trial-by-trial accuracy feedback. Although this should minimize the role of response priming, a robust accuracy NCE effect was observed. However, even without an emphasis on speeded responses, it may be that subjects automatically initiate their responses when the prime appears. The current data cannot rule out this possibility. To address this concern, Experiment 1b modified the paradigm by collecting same/different responses. With this modification of the task, the angle of the prime (e.g., "left" vs. "right") no longer indicates either response considering that the responses were "same" versus "different."

Experiment 1b: Same/Different Orientation Judgment

Introduction

Experiment 1a mapped out a transition from PCE to NCE with increasing prime duration in a nonspeeded accuracy task. However, response priming may have played a role if subjects automatically encoded the response attributes of the prime. To eliminate this alternative explanation, we conducted an otherwise identical experiment except that same/different responses were collected rather than left/right responses. Same/different judgements were initiated upon appearance of a test display that contained an outer oriented stripe and these judgments concerned the relationship between the outer stripe and the target (i.e., whether the stripe was in the same direction as the darker lines of the target display). In this task, subjects needed to identify the target in the same manner as in Experiment 1a, but they could not possibly know which button to press until the test stripe appeared. Thus, a particular orientation for the prime should not elicit a response of either "same" or "different." If the effects found in Experiment 1a reflect orientation perception (i.e., activation and habituation for the concept of left-leaning or right-leaning), then this same/different version of the task should produce results that are similar to Experiment 1a, and, furthermore, this should be the case for both "same" trials (i.e., trials where the test stripe matches the high contrast orientation of the target display) and "different" trials (i.e., trials where the test stripe matches the low contrast orientation of the target display).

Method

All methods were the same as Experiment 1a except where noted.

Participants

An equivalent power analysis as the one described in Experiment 1a suggested that a sample of 45 subjects would yield power of over 95% for the two-way interaction of interest (prime type and prime duration). The same parameters were used in the simulations with the exception of a predicted standard deviation of .15 (as the number of trials per condition in Experiment 1b was 18 as opposed to 36). To account for potential issues such as equipment malfunction, data from 58 subjects aged 18–35 were collected, and five were excluded from all analyses due to displaying accuracy below 60% across all conditions (global accuracy for the five excluded subjects was 51.3%, 55.2%, 57.2%, 58.3%, and 59.8%).

Materials

The task was identical to Experiment 1a, except that subjects were instructed to perform a same/different judgment on a gray stripe in relation to the identified target orientation (see Figure 3). The mask that followed the target display was presented by itself for 500 ms. At that point, the gray stripe appeared in the background, and subjects had to decide whether that stripe was the same orientation as the darker set of lines of the target grating. Subjects pressed the J key for "same" and the F key for "different."

Procedure

The procedure was identical to Experiment 1a, except that there were twice as many conditions considering that each trial could end with a nominally correct answer of "same" or "different," depending on the orientation of the test display stripe. Thus, when collapsing across same/different, the design was equivalent to Experiment 1a, but when considering "same" and "different" trials separately, there were 18 trials per condition per each subject, rather than 36. Pilot results indicated that the same/different task was more challenging than the forced-choice task and to accommodate this extra difficulty, the adjustment block of trials began with a contrast ratio of three rather than the contrast ratio of 2.75 used in Experiment 1a (i.e., to better ease subjects into this challenging task, they began with a slightly higher visual contrast).

The average contrast ratio of all subjects across all experimental trials was 2.7 (1.1 standard deviation).

Results

The analysis techniques were identical to Experiment 1a, except that the ANOVA included an extra factor of whether the correct response was "same" or "different."

Accuracy across conditions is displayed in Figure 6, separately for conditions in which the correct response was "same" and conditions in which the correct response was "different." Statistical analysis revealed significant main effects of prime duration, *F*(4, 208) = 6.735, *p* < .001, η_p^2 = .115; prime type, *F*(2, 104) = 14, *p* < .001, η_p^2 = .212; and whether the correct response was "same" or "different," *F*(1, 52) = 18.98, *p* < .001, η_p^2 = .267. There was a significant interaction between prime duration and prime type, *F*(8, 416) = 27.77, *p* < .001, η_p^2 = .348, and a three-way interaction between all factors, *F*(8, 416) = 2.364, *p* = .017, η_p^2 = .043.

As seen in Figure 6, the pattern between prime duration and prime type was similar to Experiment 1a, and this was true for both "same" trials and "different" trials. Reflecting a bias to respond "same," accuracy for the "same" trials was higher than for "different" trials. This accuracy difference might be the cause of the significant three-way interaction considering that the ANOVA assumes a linear scale. In other words, there may have been a ceiling effect for the "same" conditions, resulting in smaller apparent effects when measured on the bounded probability scale. To test this possibility, we conducted an ANOVA on logit transformed accuracy data, to transform probabilities (p) into log (L) likelihoods (i.e., L = log(p/(1 − p))). The results replicated the significant main effects of prime duration, *F*(4, 208) = 3.608, *p* = .007; prime type, *F*(2, 104) = 11.09, *p* < .001; and whether the correct response was "same" or "different," *F*(1, 52) = 17.73, *p* < .001. The significant interaction between prime duration and prime type was also replicated, *F*(8, 416) = 24.56, *p* < .001, but the three-way interaction between all factors was no longer significant, *F*(8, 416) = 1.357, *p* = .214, supporting the claim that the apparent three-way interaction was indeed an artifact of the nonlinear probability scale.

Uncorrected posthoc pairwise t-tests (see Table 2) revealed that the shift from priming benefits to deficits happened slightly faster than was the case for Experiment 1a, with no significant difference between the compatible and incompatible means at 34 ms for both "same" and "different" conditions (in Experiment 1a, priming benefits were still present at this point).

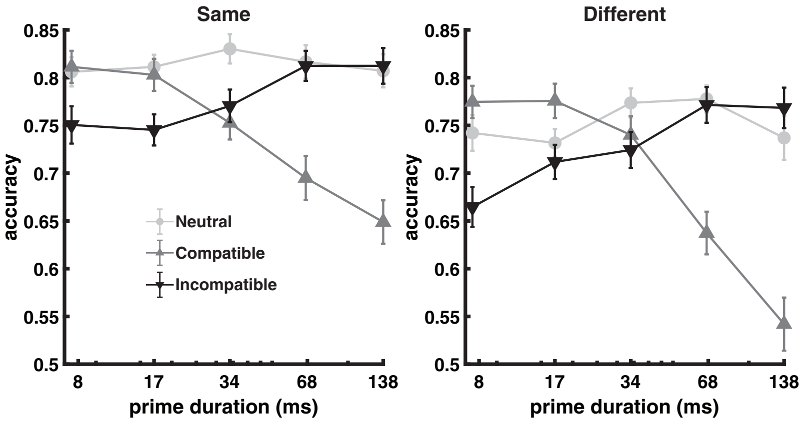
Discussion

All of the key priming compatibility effects found in Experiment 1a were replicated in Experiment 1b, except for a slightly faster transition from positive to negative priming in Experiment 1b. Critically, the priming pattern was qualitatively the same for both "same" trials and "different" trials except as modulated by higher accuracy for "same" trials (i.e., there was an overall bias to response "same"). The slightly faster transition to negative priming may reflect the higher on-average contrast ratio for Experiment 1b as compared with Experiment 1a (2.7 vs. 2.3). In other words, the darker lines of the target orientation were on-average darker for Experiment 1b than for Experiment 1a. The observation that greater target information results in a faster transition to negative priming is consistent with prior studies of word and face priming, which found that increasing target duration can shift the priming crossover pattern to favor stronger/faster negative priming (Huber, Shiffrin, Lyle, et al., 2002; Rieth & Huber, 2010; Weidemann et al., 2008).

Because subjects could not possibly know the orientation of the test stripe at the time when the prime appeared, these results rule out an alternative explanation for the finding of Experiment 1a in terms of response priming. Consistent with the hypothesis that

Figure 6

Average Accuracy for Experiment 1b as a Function of Prime Duration (Log Scale) and Priming Condition, Broken Down by Whether the Correct Answer in a Trial Was "Same" or "Different"



Note. Neutral primes consisted of vertical lines displayed at 90°. Compatible refers to primes that matched the target angle (45° or 135°), and incompatible refers to primes that did not match the target angle and instead matched the foil angle (135° or 45°). Error bars are plus and minus one standard error of the mean (2 *SEM*s in total).

Table 2

Paired t-Tests Comparing Accuracy in the "Compatible" Condition Against Accuracy in "Incompatible" Condition for Same and Different Conditions Separately in Experiment 1b (df = 52)

Prime duration	8 ms	17 ms	34 ms	68 ms	138 ms
			Same condition		
t-statistic	2.38	2.57	−0.75	−4.29	−5.05
p-value	$p = .021$*	$p = .013$*	$p = .45$	$p < .0001$***	$p < .0001$***
Mean diff.	0.061	0.058	−0.018	−0.117	−0.164
			Different condition		
t-statistic	4.65	2.77	0.599	−4.73	−6.79
p-value	$p < .0001$***	$p = .0077$**	$p = .55$	$p < .0001$***	$p < .0001$***
Mean diff.	0.110	0.064	0.016	−0.134	−0.226

* $p < .05$. ** $p < .01$. *** $p < .001$.

these effects reflect perception rather than response, the same priming crossover pattern was observed regardless of whether the correct response was "same" or "different."

Experiment 1c: Neural Habituation Model Fitting

Introduction

The neural habituation model was only fit to Experiment 1a because that experiment specifies the dynamics of orientation perception without needing to model decisional processes, such as the bias to respond "same" that was found in Experiment 1b. There were two goals of this model-fitting exercise: (a) can the model fully explain the data with relatively few free parameters (i.e., is the model quantitatively sufficient); and (b) what are the best-fitting perceptual dynamics, which can then be applied to the NCE literature. In Experiment 2 these perceptual dynamics were fixed, and the model was augmented with a response layer in applying the model to the NCE literature. Thus, the simulations reported in Experiment 2 used parameter values that resulted from fitting Experiment 1a.

The neural habituation model represents perception in a hierarchical organization that is similar to the ventral visual stream, with each progressive layer of the model representing more complex visual stimuli. As the model can be applied to different tasks using different stimuli, the specific structure of the model will vary accordingly. In all versions of the model, the first layer represents location-specific retinotopic information in primary visual cortex; this layer then feeds into a second, higher-level layer, but the representation and dynamics of the second layer will depend on the stimuli (e.g., letters for word identification or face-features for face identification).

In the present study, the second layer represents orientation perception regardless of visual field position or spatial frequency (i.e., a more generalized concept of orientation). Given the simple nature of the stimuli and task used in this experiment, no further layers were needed; for experiments using more complex visual objects additional perceptual layers are needed (i.e., a word priming task would need a third layer representing whole words, with that layer receiving input from letter perception in the second layer), and for experiments involving comparisons between a current test stimulus and something in the past (e.g., same/different decisions or episodic familiarity), an additional memory layer is needed, for instance capturing long-term memory (Huber, Clark, et al., 2008) or working memory (Jacob & Huber, 2020; Rusconi & Huber, 2018).

The model implements neural habituation with a series of mathematical equations as described in detail in the Methods section below. These equations calculate dynamically varying properties every millisecond for idealized rate-coded neurons (a so-called "node"), with each node capturing the behavior of a large assembly of spiking neurons that have similar inputs and outputs. These dynamically varying properties include membrane potential, which is related to average firing rate of the assembly, and the current level of synaptic resources (e.g., available neurotransmitter), which formalizes neural habituation. The output of each node is the product of these two variables (i.e., the product of average firing rate and the effect of each action potential).

Because pyramidal cells are largely the same throughout the cerebral cortex (e.g., the threshold membrane potential for producing an action potential is the same for all spiking neurons), most of the parameters used in the present version of the model are identical to those used in prior publications. Only three free parameters were fit to the present results (a noise parameter and two temporal integration parameters, one for each layer).

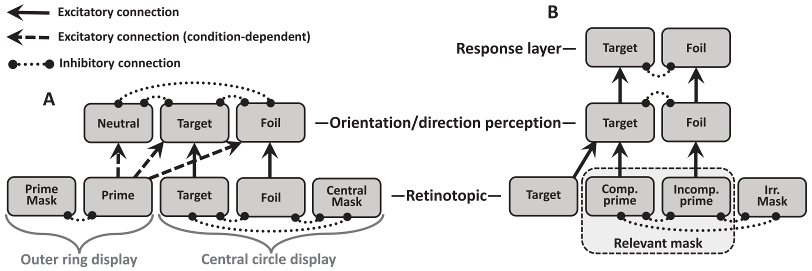
Method

The model structure used in this experiment is detailed in Figure 7A. Two layers were used: a retinotopic layer that consists of visual nodes divided into two groups, and an orientation perception layer that consists of three nodes. The two groups in the retinotopic layer correspond to the two task relevant screen positions: the outer ring (where prime and prime mask were displayed) and the central circle (where the target plaid and subsequent target mask were displayed). Within each group, all retinotopic nodes inhibit each other, capturing the winner-take-all effect of local inhibitory interneurons (Carandini & Heeger, 1994).

The retinotopic layer feeds into the orientation perception layer. The prime node is mapped to one node in this layer depending on the condition being simulated (neutral; compatible, which maps to the target node; or incompatible, which maps to the foil node). Note that while the model nodes are labeled as "target" and "foil," this labeling is simply for convenience; these nodes correspond to one orientation or the other (45° or 135°), with the target node being the orientation that was the same as the prime for compatible trials and foil node being the orientation that was the same as the prime for incompatible trials. The system does not know which

Figure 7

Habituation Model Structures



Note. (A) structure used to capture accuracy for the forced-choice orientation priming experiment (Experiment 1a, with model simulation reported in Experiment 1c). (B) Structure used to generate reaction time predictions for arrow priming NCE experiments (Experiment 2). Each gray box is a single idealized "node" that contains a variable for average firing rate and a second variable for the current level of available synaptic resources. The multiplication of these two variables determines the output of the node at each simulated millisecond.

orientation is the target and instead a guess is made based on the relative activations of the two orientations contained in the overlapping grating that create the plaid display.

The present implementation of the model did not include any longer-term learning or intertrial effects and all effects were assumed to arise from the within-trial perceptual dynamics. However, full recovery from habituation can take several seconds, producing effects from the last trial, or over the course of several trials. For the current paradigm, the effects of previous trials washout on average, considering that the trial sequence was randomized. Nevertheless, the neural habituation model has successfully explained across trials effects, such as repeating a word from one trial to the next in a change detection task (Davelaar et al., 2011) or repeating a word up to ten times across twenty trials in a semantic judgment task (Tian & Huber, 2010, 2013). In addition, the model has been applied to short-term priming in a long-term recognition task (Huber, Clark, et al., 2008), using a familiarity layer whose weights were modified as a function of prior study.

Model input to the retinotopic layer nodes was all-or-none (zeros or ones, depending on condition and time point within the simulation), with the exception of input to the foil node, which was set to .25 during times when the target plaid was displayed. This lower value of .25 captures the visual contrast difference between the high contrast target orientation (input of 1.0) and the lower contrast foil orientation (input of .25). The value of .25 is somewhat arbitrary and does not necessarily correspond to a contrast ratio of four considering that visual contrast is not necessarily the same thing as input to primary visual cortex (V1) from the lateral geniculate nucleus (Michelson contrast of an oriented grating is more likely related to the output of simple cells in V1 rather than input to these cells). This value of .25 was determined from explorations with the model prior to data collection.

The sequence of inputs to the model followed directly from the times displayed in Figure 3. While the prime was presented, the retinotopic prime node received input of 1.0. When the target display grating was presented, the prime node's input became .0, the target and prime mask nodes received an input of 1.0, and the foil node received an input of .25. Finally, once the central mask was presented, its node input became 1.0, the target and foil node inputs became .0, and the prime mask node remained at 1.0.

The activity of each node was captured with two dynamically varying terms, with the product of these determining the output that the node can provide to other nodes. The first term is membrane potential (v), which is compared with the fixed firing threshold (θ) to determine the probability of an action potential (i.e., firing rate). Because the node implements the activity of many neurons, simulations use this firing rate rather than simulating spiking behavior. However, an action potential does not necessarily produce a postsynaptic response if there are no neurotransmitter vesicles available to release, and so the second term captures the current level of neurotransmitter resources (a). Equation 1 is the product of the firing rate and neurotransmitter resources, which determines the output of the node (o). If membrane potential is below the firing threshold, the output is zero.

$$o = (v - \theta)a \qquad \text{(Equation 1)}$$

In simulations, these terms are updated every millisecond. At the start of the simulation, output and membrane potential are set to 0, and neurotransmitter resources are set to 1; these terms are bounded between 0 and 1 (i.e., if the update equation would result in them going out of bounds, they are set to the bound). Membrane potential (v) is updated according to Equation 2, which computes Δv for each node $i$in each layer n. The first bracketed term corresponds to excitatory inputs (bottom-up connections from the $n-1$layer) modulated by connection weight w. In the present experiment, the weights are set either to 1 or 0 according to whether two nodes are connected. The second bracketed term corresponds to inhibitory inputs, which are a combination of constant leak (L) and lateral inhibition (modulated by inhibition strength I), generated by mutual inhibition between the nodes of a layer or group within a layer (and thus affected by their present level of activity). The level of lateral inhibition is the summation of all nodes within the layer (or group), capturing the effect of all-to-all connected inhibitory interneurons, which serve to limit excessive activity. This all-to-all summation includes self-inhibition (not shown in Figure 7). Finally, *S*corresponds to the rate of integration, also unique to each layer.

$$\frac{\Delta v_i^n(t)}{S_n} = \left(1 - v_i^n\right)\left\{\sum_{\forall j} w_{ij} o_j^{n-1}\right\} - v_i^n\left\{L + I\sum_{\forall j} o_l^n\right\} \qquad \text{(Equation 2)}$$

The amount of neurotransmitter resources (a) within a node is updated according to Equation 3, which computes Δa as a function of neurotransmitter depletion rate (D) and recovery rate (R), as well as the node's output (o) and its layer's rate of integration (S).

$$\frac{\Delta a}{S} = R(1 - \ a) - D \times o \qquad \text{(Equation 3)}$$

These equations are the same as in all prior publications reporting simulations with the habituation model.

Accuracy predictions are obtained from the orientation perception layer. The relevant measure is the difference in activation (m) between the orientation that is most active and the alternative orientation at the time when the "winning" orientation reaches its maximum response. For the best-fitting parameter values, the winning orientation was always the target orientation, although this need not be the case in principle (e.g., even longer prime durations might have driven accuracy below 50%, corresponding to a situation where the foil orientation was more active). Simulations with the model are deterministic, and in mapping model behavior onto the accuracy scale, we captured perceptual noise by adding mean zero normally distributed noise with a standard deviation of *N*. Thus, although the average difference was m, sometimes the foil orientation could be perceived to be more active than the target orientation because of perceptual noise, resulting in an error. Accuracy is therefore the probability that the activation difference is positive, rather than negative, which is simply the inverse cumulative normal distribution (Φ) with a mean of *m* and standard deviation *N* (Equation 4).

$$P(correct) = 1 - \ \Phi(\mu = m, \ \sigma = N) \qquad \text{(Equation 4)}$$

As previously described, three parameters were fit to the data: the noise term *N*, and the rate of integration of both layers (S_1 and S_2). They were fit to the average accuracy for each of the 15 experimental conditions (five prime durations for each of the three priming conditions), minimizing the binomial likelihood ratio test statistic G2, which is distributed as a $\chi 2$ (Riefer & Batchelder, 1988).

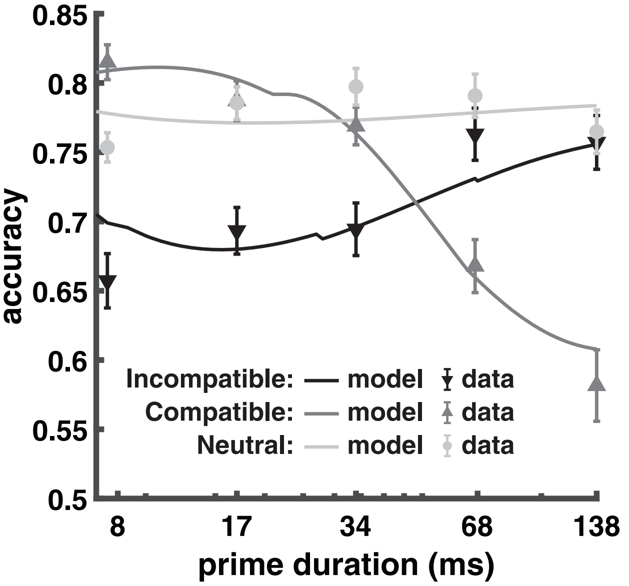
The other parameters were inherited from application of the model to word priming results of Rieth and Huber (2017): $\theta = .15$, $R = .022$, $L = .15$, $I = .9844$, $D = .324$, although we note that these same parameter values have been used in nearly all prior applications of the habituation model.

Results

Fitting the accuracy data (15 conditions with three parameters) yielded the following parameters: $S_1 = 0.0756$, $S_2 = 0.1918$, and $N = 0.3248$, with G2 = 51.16 and 90.1% of the variance accounted for. In creating Figure 8, these parameter values were then used in simulations to generate accuracy predictions not just for the tested prime durations, but for every prime duration in steps of 1 ms from the shortest prime duration of 8 ms to the longest prime duration of 138 ms.

Figure 8

Model Accuracy With Best-Fitting Parameters (Solid Lines) Versus Observed Data (Points With Error Bars) Across Prime Duration (Log Scale) for the Results of Experiment 1a



Note. Although the model is deterministic, in a few instances, model accuracy appears to produce a small discrete change as a function of prime duration. This reflects the use of a discrete update equation implemented at each millisecond. Accuracy is determined by the difference in output activation between the target and foil orientation nodes at the time when the target node reaches its maximum value, and these small discrete steps reflect a change of one millisecond in terms of which time point after presentation of the target plaid produced the largest response for the target node.

With best-fitting parameters, the model captured the key interactions of prime type and prime duration, with neutral primes generating flat accuracy across prime durations, compatible primes shifting from performance benefits to deficits, and incompatible primes shifting from performance deficits to baseline performance. There are a few instances where model behavior falls outside of the error bars of the observed data, but it is important to keep in mind that this is a highly constrained fit; the model necessarily produces a rise-then-fall trend in terms of increasing and then decreasing positive priming with increasing prime duration, with the parameter values only serving to dictate the rapidity of this trend (captured by the value of S1), whether the trend achieves strong negative priming (captured by the value of S2), and a monotonic rescaling of model behavior to the accuracy scale (capture by the value of *N*). There is a slight n-shaped curve as a function of prime duration for the neutral condition whereas the model is either flat or u-shaped, reflecting the rise and fall of interference from the horizontal prime orientation. If reliable, this n-shaped curve may reflect an attentional factor beyond the scope of this relatively simple model (e.g., performance for all of the 8-ms prime conditions might be lower than expected owing the spatial distraction of the briefly flashed prime).

As described in the Experiment 1a Discussion, we hypothesized that priming effects are due to the amount of lingering activation for the prime (positive priming) as compared to the amount of lingering habituation from the prime (negative priming), with both of the factors carrying over, affecting the response to the target display. To illustrate model behavior, we plot the time course of the output variable and synaptic resources variable of the two nodes in the orientation perception layer for the shortest and the longest prime durations separately for the compatible and incompatible prime conditions (see Figure 9). This figure shows the difference between target and foil output (the vertical arrow in each condition) at the moment when the target output reaches its maximum value. This is the measure used to generate accuracy predictions; the larger this measure, the higher the predicted accuracy. As seen in the figure, a 138-ms compatible prime (upper right graph) resulted in low accuracy (shorter vertical arrow) due to a relatively small response to the target display, which was caused by habituation for the target orientation (dashed black line) that accrued over the time course of the prime presentation. This negative priming can be contrasted with the low accuracy in the incompatible 8-ms condition (lower left graph), where the target response was robust, but lingering activation from the prime, which matched the foil, resulted in a smaller difference (shorter arrow) between the target and foil orientations.

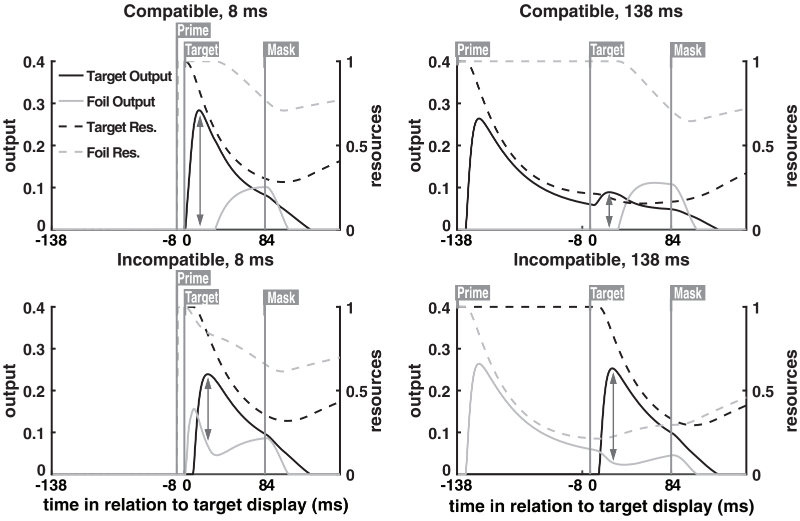
Discussion

The model successfully captured the key priming effects, demonstrating that neural habituation can explain accuracy for this orientation identification task. The constrained nature of the parameter fits (three free parameters), with most of the parameter values inherited from prior publications that used different tasks, highlights the model's ability to generalize to new tasks and stimuli. The success of this generalization supports the hypothesis that neural habituation is a general mechanism for visually parsing the perceptual response to the current stimulus from recently viewed stimuli.

According to the model, priming effects reflect both the level of activation and the level of resource depletion generated by the prime display, which carry over and affect the response to the target display. Short duration primes cause the neural representation of the primed orientation (whether compatible or incompatible) to

Figure 9

Output (Solid Lines) and Synaptic Resources (Dashed Lines) for the Target Node and the Foil Node in the Orientation Perception Layer for Four Different Conditions Based on a Short (8 ms) or Long (138 ms) Duration Prime That Matched The Target (Compatible) or Foil (Incompatible) Orientation



Note. The gray vertical lines with labels indicate when the corresponding stimulus appeared in the display sequence. The arrows show the difference in output between the target and foil orientations at the time when the target orientation reached its maximum output following presentation of the target display. This difference determined accuracy according to Equation 4. These graphs illustrate how the model produces a transition from positive to negative priming with increasing prime duration: Following an 8-ms prime, lingering activation from the prime in the incompatible condition (solid gray line in the lower left graph) results in a smaller difference between the target and foil response (vertical arrow in the lower left graph as compared with the upper left graph) whereas following a 138-ms prime, lingering habituation from the prime in the compatible condition (dashed black line in the upper right graph) results in a smaller difference between the target and foil response (vertical arrow in the upper right graph as compared to the lower right graph).

become active, but due to the short duration, few synaptic resources are spent, and no significant amount of neural habituation is generated. Once the prime is replaced by the target display, this lingering activation from the prime affects the relative levels of activation for the target and foil orientations contained in the target display. If the prime is compatible with the target, this lingering activation makes target identification easier (i.e., larger difference between target and foil). On the other hand, if the prime was incompatible, this lingering activation favors the wrong answer. In summary, lingering activation from the prime produces positive priming, boosting the response of the matching orientation.

This situation changes as prime duration increases: The longer that the prime appears, the more synaptic resources are spent by neurons that preferentially respond to the prime orientation. The effect of this short-term synaptic depression is two-fold: First, it reduces the amount of lingering activation for the prime's orientation (thus reducing the source of positive priming), and second, it increases the amount of lingering habituation for the prime's orientation, making it difficult to reactivate that orientation in response to the target display. This lingering neural habituation can produce negative priming rather than just a reduction in positive priming. This can be thought of as a kind of repetition blindness (Kanwisher, 1987) for the target orientation due to lingering habituation for the orientation of the target; although the retinotopic cells corresponding to the target's orientation provide an adequate response to the target display, they fail to drive the more generalized perception of the target's orientation. In the case of incompatible primes, this repetition blindness is beneficial because it serves to reduce the orientation response for the incorrect foil orientation (the dip in the light gray line at the time of the target display following an incompatible 138 ms prime, as seen in the lower right graph of Figure 9). Because lingering neural habituation offsets lingering activation for the foil orientation, the target orientation is identified as easily as in the baseline condition that did not include priming.

Next, we ascertain whether these same orientation perceptual dynamics (i.e., the neural habituation model with parameter values fixed according to this fit of Experiment 1a) can explain reaction times in the NCE literature.

Experiment 2: Neural Habituation Account of the NCE

Introduction

Over 2 decades have passed since the NCE was first reported (Eimer & Schlaghecken, 1998), with subsequent studies producing varied results that seem to contradict each other (i.e., one study produces an NCE, while a different study with slightly different timing or masks produces a PCE). While there are several possible reasons for the large variability among the results, such as differences in stimuli and screen positions, we focus on the timing of the display sequence. We hypothesize that this is the primary factor underlying the conflicting effects, and we support this hypothesis by applying the neural habituation model to a typical NCE paradigm, showing the manner in which other factors such as the type of mask can interact with timing in the display sequence.

When using a representative paradigm such as the one in Figure 2, in which masked primes do not share the same screen position as target, prior studies found PCE for irrelevant masks (Lleras & Enns, 2005, 2006) and no masks (Klauer & Dittrich, 2010) when using shorter prime-mask ISIs, but NCE for longer ISIs (Klauer & Dittrich, 2010). Additionally, prime duration appears to have a similar effect as ISI, with longer duration primes producing an NCE for irrelevant masks (Jaskowski, 2008; Klapp, 2005) as compared with equivalent studies that used shorter duration primes, which instead found PCE (Lleras & Enns, 2004, 2005, 2006). In contrast to these varied PCE/NCE results with no masks or irrelevant masks, when a relevant mask is used, an NCE is consistently observed (Eimer & Schlaghecken, 1998; Jaskowski, 2008; Klapp, 2005; Lleras & Enns, 2004; Praamstra & Seiss, 2005).

We examined the predictions of the neural habituation model for these manipulations using two representative prime durations, one shorter (16 ms) and one longer (40 ms), and several ISIs from 0 ms (mask presented immediately after prime) to 160 ms. For each of these situations, we simulated conditions with relevant masks, irrelevant masks, and no masks. This application of the habituation model used the same perceptual layers as presented in Experiment 1c, and the same parameter values for these layers as determined by the fit to Experiment 1a.

To capture reaction times, we added a response layer to the model (see Figure 7B) that did not include neural habituation, with this layer accumulating response information at a slower pace across the entire display sequence. The lack of habituation and a slow integration rate reflect the function of this layer, which is to accumulate evidence, rather than temporal parsing of the RSVP display sequence. This accumulation of response information is akin to the evidence accumulation decision processes contained in sequential sampling RT models (Smith et al., 2004), although evidence accumulation in this layer is implemented with competition between leaky neurons, making this layer an instantiation of the leaky accumulator RT model of Usher and McClelland (2001).

This response layer does not produce perceptual priming effects, but rather accumulates the output of perceptual priming effects during the display sequence to determine how much information favors the correct answer. Thus, unlike the model as applied to the Experiment 1a, where accuracy reflected the relative activation of target versus foil in the orientation perception layer, RT in a typical NCE study is assumed to reflect the totality of target information accumulated across the display sequence. This target information will always support the correct answer considering that the target remains onscreen until a response is given, but the question of interest is how quickly this information accumulates.

Model Structure and Retinotopic Input

The model structure for the NCE paradigm is shown in Figure 7B. The retinotopic layer consisted of four nodes: a target node (visual features of the target), a compatible prime node (visual features of a prime that matches the direction of the target), an incompatible prime node (visual features of a prime that matches the other possible response direction), and an irrelevant mask node (visual features of a mask with vertical and horizontal lines). Relevant masks were simulated by activating the compatible prime node and the incompatible prime node at the same time, considering that relevant masks were composed of superimposed arrows pointing to the left and to the right (see Figure 2) and are thus formed by overlaying the two prime types.

In the retinotopic layer, both prime nodes (which, together, form the relevant mask) and the irrelevant mask node inhibit each other due to their shared screen position, but they do not inhibit the target node considering that the target is the only stimulus presented in the flanking screen positions. The inhibition originating from the irrelevant mask node was multiplied by two considering that the irrelevant mask node consisted of both horizontal and vertical lines. This makes the irrelevant mask comparable with the relevant mask in terms of its masking potential (i.e., the summation of arrows pointing both directions was assumed to be as effective a mask as the summation of horizontal and vertical lines). Alternatively, we could have simulated the irrelevant mask condition by having both a horizontal and vertical orientation, with both of these activated by the irrelevant mask, which would have produced the same result as this multiplication by two assumption.

The retinotopic layer feeds into a direction perception layer that is equivalent to the orientation perception layer from Experiment 1c. The two nodes in the direction perception layer inhibit each other through lateral inhibition. Finally, the direction perception layer feeds into the response layer.

Over the course of a simulated trial, input to the appropriate retinotopic nodes was set to 0 or 1 according to the condition that is simulated, with the timings appropriate to that condition. In the case of the no mask condition, all inputs were set to 0 during the time the mask would typically be presented (in other words, all inputs were set to 0 both during the prime-mask ISI duration and the “mask” duration).

Response Layer: RT Predictions

The response layer has the same structure as the direction perception layer, but with different parameter values, reflecting the assumed function of this layer, which is to accumulate response information across the display sequence.

The neural habituation model is deterministic (i.e., if the model were run twice on the same trial, it would exhibit identical behavior each time). However, a key element when explaining reaction times is capturing the right-skewed response distribution and the way the shape of the RT curve changes with changes in response bias. In mapping model behavior onto RT, we assume that the drift rate of a noisy evidence accumulation decision process is proportional to the maximum output of the target node in the response layer (i.e., a stronger target response produces faster drift toward the correct answer). The simplest form of a single answer diffusion model is a Weiner process (Usher et al., 2002), which is described by a reparameterized inverse Gaussian or Wald distribution, with one parameter representing the drift rate (i.e., rate of evidence accumulation) while a second parameter represents the decision boundary (i.e., the amount of evidence that must be accumulated before a decision is made, thus capturing response bias). This assumption is similar to the diffusion race model that Potter et al. (2018) used to explain same/different RT distributions for a word priming experiment, except that in the current case there is just one racer (the correct target direction), that is guaranteed to win (100% accuracy), with the diffusion process capturing the time it takes to reach the decision boundary. The same decision boundary was used for all conditions (all mask types and durations).

Prior NCE studies only report average RT and so the current application to the NCE literature only considers average RT. This is easily achieved because the average of the reparameterized inverse Gaussian is equal to the drift rate divided by the decision boundary, and thus a prediction for average RT is simply the maximum output of the target response node divided by the decision boundary parameter. However, we note that these assumptions make predictions about the shapes of the RT distributions in different conditions, predictions that await future study.

The response layer inherited the same model parameters as reported in Experiment 1c and in prior publications, except that the depletion value $D$was set to zero, as the function of this layer is accumulation of response information rather than temporal parsing of the display sequence. For the response layer nodes, the rate of integration $S$and the voltage leak value Lwere free parameters, considering that response information might accumulate more slowly (small S) and the rate at which information dissipates (L) might need to be adjusted to modulate the degree to which response information from the prime and relevant mask carry over into the decision process. Thus, the rate of integration and the leak value, along with the decision boundary mentioned above, were fit to representative literature results. The values for these three parameters were 7.71 (boundary height), .0141 (rate of integration), and .302 (voltage leak).

The literature results used to optimize these response layer parameters were selected from paradigms consistent with the perceptual node arrangement shown in Figure 7B. More specifically, we only considered studies in which the target did not share a screen position with the masked primes. Of the experiments that met this requirement, a representative set was chosen for their use of different timings and their investigation of different mask types. The experiments used were: (a) the prime/mask-only-at-fixation condition of Lleras and Enns (2005), which found NCE for relevant masks and PCE for irrelevant masks, and used a 15 ms prime and no prime-mask ISI; (b) Experiment 1, Group C, and Experiment 2, Group C, of Jaskowski (2008), which found NCE for both relevant and irrelevant masks, with the NCE for relevant mask being stronger, and the use of a 25 ms prime and 25 ms or 75 ms ISI; and (c) Experiment 5 of Klauer and Dittrich (2010), which found PCE for no masks with a prime-target ISI of 120 ms, and NCE with a prime-target ISI of 240 ms, with prime duration being 40 ms. Pooling results from different experiments was necessary due to the prevalence of conflicting results in the NCE literature, as outlined in the Experiment 2 Introduction, and the lack of within-subjects parametric studies that examine a wide range of prime durations, prime-mask ISI, and mask type.

Given that different experiments used different procedures, different instructions, different subject populations, and different stimuli (e.g., Klauer & Dittrich, 2010 used up- and down-facing arrows instead of left and right), we did not expect a close fit to the RT data (our fit to the experiments listed above resulted in 70.9% of the variance accounted for). Many of these studies only reported RT priming effects (i.e., difference between the compatible and incompatible conditions), and so we did not attempt to fit the separate RTs of each condition, instead focusing on this RT difference measure (although we note that the model makes predictions regarding RT for each condition separately). The goal of this optimization was to determine if decision parameters could be found that account for the qualitative trends in the literature when using perceptual parameters that were fixed according to the application of the model to orientation priming accuracy results (Experiment 1a).

Results

Simulation results for two different prime durations are shown in Figure 10. Average RT for the compatible and incompatible conditions were compared, with the direction of this subtraction such that a positive value is a PCE (faster responses when the prime matches the target) and a negative value is a NCE (slower responses when the prime matches the target).

As typically found in the arrow priming literature, the use of relevant masks produced a NCE regardless of the prime-mask ISI, with NCE magnitude increasing as the ISI increased, until it plateaued around 100 ms for both prime durations. Both irrelevant masks and no masks (blank screen presented instead of a mask) generated PCE at shorter ISIs, and these PCEs decreased as ISI increased, eventually becoming NCEs. Additionally, this shift from PCE to NCE occurred for a shorter ISI when the prime was presented for a longer duration.

Discussion

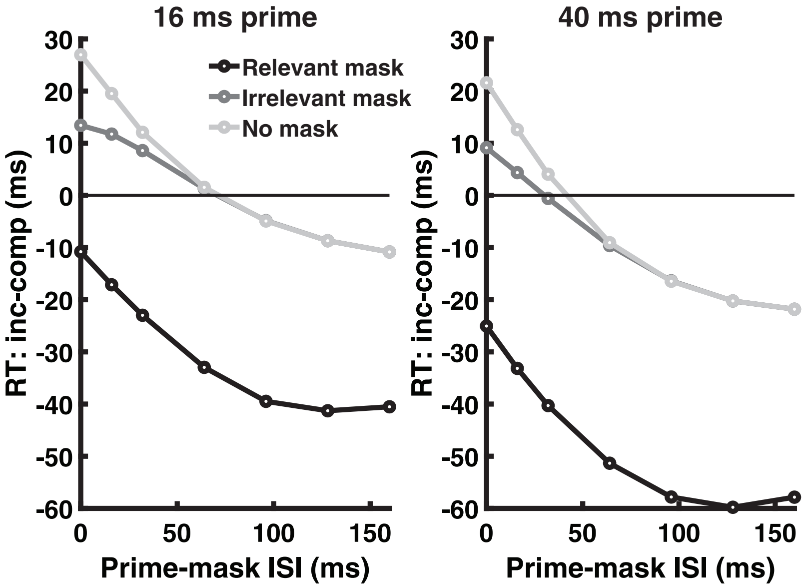
Using perceptual dynamics determined by the orientation priming task with threshold accuracy (Experiment 1a), the neural habituation model captured the various RT priming effects reported in the NCE literature. The key experimental factor was prime-mask ISI; increases in ISI changed the priming effect from PCE to NCE. However, overlaid on this ISI effect was a large shift toward NCE with relevant masks and a moderate shift toward NCE when prime duration was increased. Finally, there was a small shift toward stronger PCE for no mask compared with irrelevant masks, but this was only for short ISIs.

According to the neural habituation model, the prime-mask ISI and prime duration have similar effects, with increases in either or both of these experimental factors producing greater habituation for the perceptual direction of the prime. The perceptual system is slow to recover from this habituation, and lingering habituation in the direction perception layer carries over to the time when the target appears. In the case of a compatible prime, this lingering habituation makes the perceptual system slow to perceive the direction of the target, producing an NCE effect.

Why is there a PCE effect for some conditions according to the model? Key here is the role of the response layer, which slowly accumulates subthreshold response information (in the form of membrane potential) across the display sequence. Because this layer is relatively slow, and because this layer does not habituate, the response direction of the prime is added onto the total accumulated response information. In other words, because the decision system is accumulating response information throughout the trial, response priming can produce a PCE. This is what occurred with a shorter ISI in the no mask or irrelevant mask conditions. In these cases, the detrimental effect of lingering perceptual habituation is not sufficient to overcome this positive response priming effect. However, for longer ISIs, perceptual habituation becomes the

Figure 10

Neural Habituation Model Simulation Results Comparing Two Prime Durations (Left Graph With 16 ms Prime and Right Graph With 40 ms Prime) and Prime-Mask ISIs Ranging From 0 to 160



Note. The *y*-axis shows the average RT for the incompatible prime condition minus the average RT for the compatible prime condition. Values above 0 represent a positive compatibility effect; values below 0 represent a negative compatibility effect. Relevant masks were the superposition of the compatible and incompatible primes and irrelevant masks were the superposition of horizontal and vertical lines.

stronger factor, producing NCE despite response priming. Thus, the response layer is crucial for a full explanation of the NCE literature, but of note, this response priming explains PCEs rather than NCEs. In the absence of the response layer and using the same parameters from Experiment 1a, the perceptual layers of the model always produce a perceptual NCE effect (unlike Experiment 1a) because the imposition of the mask between prime and target eliminates any lingering perceptual activation (i.e., the intervening mask eliminates the source of positive perceptual priming), leaving only perceptual habituation. In summary, response priming can produce PCE when examining RT to easily seen targets, but sufficient perceptual habituation can slow down perception of the target (e.g., repetition blindness for the target) and offset this response priming effect.

Why is there no PCE for relevant masks according to the model? Because the mask is the superposition of both possible response directions, it activates both directions in the direction perception layer, providing additional response information to the response layer favoring both responses. When a compatible prime is presented, some response information accumulates for the correct answer, but when the relevant mask appears, response information accumulates for the wrong answer as well. Furthermore, because the retinotopic and perceptual direction activations favoring the target direction are habituated by the time that the relevant mask appears, perception of the relevant mask will favor the incorrect direction (as they are in competition due to mutual inhibition). Thus, following a compatible prime, the mask primarily provides response priming for the incorrect direction. In brief, the new aspects of the mask are perceptually salient, and so the direction other than the prime “pops out” from the mask. This is the core idea behind the object updating explanation of the NCE (Lleras & Enns, 2004), and the perceptual habituation model can be viewed as a formal implementation of this idea. However, unlike the object updating model, the neural habituation model can produce NCE even for irrelevant masks or no masks, provided that the ISI is sufficiently long as to allow additional habituation. As considered in more detail in the General Discussion below, although the neural habituation account of the NCE literature is fundamentally rooted in perceptual dynamics, it nevertheless contains response priming, similar to the self-inhibition account (Schlaghecken et al., 2009).

Finally, the small difference between the no mask and irrelevant mask conditions is a masking effect (i.e., irrelevant masks are more effective when presented immediately after the prime, rather than at a delay). With zero ISI, the irrelevant mask inhibits perception of the prime at both the retinotopic and perceptual direction layers, and because the prime’s orientation is more weakly perceived, there is less response priming accumulated from the prime. By increasing ISI, the irrelevant mask loses its ability to reduce the response priming effect.

General Discussion

The neural habituation model successfully explained behavior in the orientation identification accuracy task (Experiment 1) and used these perceptual dynamics to explain key results from the arrow direction RT NCE paradigm (Experiment 2). The model inherited most of its parameter values from a word priming experiment by Rieth and Huber (2017), demonstrating its ability to generalize across tasks and stimuli. This generalization provides additional support for the theory that neural habituation serves a vital function in perception, parsing the stream of visual objects by habituating to recently viewed objects, and thus allowing unobstructed perception of subsequent objects. However, this mechanism comes at a cost, producing repetition deficits. Neural habituation exists at all perceptual levels, producing repetition deficits for repeats of the same stimulus in the same location, but also higher-level repetition deficits. In prior work, the neural habituation model explained higher-level repetition deficits such as semantic satiation to a repeated word (Tian & Huber, 2010) or the failure to perceive that a second target belonged to the target category in the attentional blink task (Rusconi & Huber, 2018). In the current study, the neural habituation model explained higher level repetition deficits for the orientation or direction of lines and arrows, regardless of where these visual objects appeared on the display screen.

When the NCE was first discovered, it was thought to reflect automatic motor response inhibition. This was based on the observation that ERPs showed motor preparation favoring the response implied by the prime when it was presented, which rapidly switched to one favoring the alternative answer when the subsequent mask appeared (Eimer & Schlaghecken, 1998). Casting doubt on this explanation, other studies found that that the NCE switched to a PCE when an irrelevant mask was used, rather than one consisting of overlapping arrows (Lleras & Enns, 2004, 2005, 2006). This supported a perceptual explanation, but additional studies cast doubt on a perceptual explanation, revealing that longer duration primes or a longer duration blank screen before an irrelevant mask recovered the NCE (Jaskowski, 2008; Klapp, 2005). In the absence of a resolution to this theoretical debate, recent studies reverted to the motor response inhibition explanation, using the NCE to study motor disorders (Rawji et al., 2020; Stenner et al., 2018). However, if the response inhibition explanation is incorrect, the conclusions drawn from these studies may be misleading. We successfully modeled the major results of the NCE literature, capturing situations that cause the NCE to switch to a PCE, or vice versa. Our account is perceptual at its core, and neural habituation in higher-level perceptual representations (i.e., a cognitive aftereffect) is the primary explanation of the NCE. However, response priming plays a role, but is facilitatory rather than detrimental, explaining situations in which a PCE is observed instead of an NCE.

Separate Causes of Positive and Negative Repetition Effects

When a target stimulus immediately follows the prime, the model explained positive or negative repetition effects as reflecting tradeoffs between two perceptual factors, which operate at different time scales, resulting in a transition from positive to negative priming with increasing prime duration. For each simulated node in the model, neural activation (average firing rate) and habituation (synaptic depletion) are separate variables, with the product of these determining output. Lingering activation is the factor behind positive repetition effects and lingering habituation is the factor behind negative repetition effects. Because habituation is driven by prior output, it lags, and is slow to recover; after a prior bout of activity, neural representations remain habituated for a period of time, making it difficult to reactivate the same neural representations when a repetition occurs. This sluggish response can produce a failure to perceive the direction of the target, as in the case of Experiment 1, or produce a relatively slow response to the target, as in the case of the NCE literature addressed in Experiment 2.

Although habituation explained negative priming effects for both the orientation priming task in Experiment 1 and the NCE literature in Experiment 2, the basis of the positive priming in each case was different. For the orientation task (Experiments 1a and 1b), the target was presented immediately after the prime, with no interleaving mask. Therefore, at the time when the target was perceived, the prime was still active, and the beneficial effect of lingering perceptual activation produced positive priming, provided that the prime duration was too short to produce strong habituation. In contrast, the arrow priming task in the NCE literature presented a mask between prime and target, which eliminated perceptual activation for the prime. Thus, there was no positive priming effect from lingering perceptual activation. However, in applying the model to the RT decision process, a slow nonhabituating evidence accumulation layer was added to the model to capture the decision process. This layer accumulates information across the entire trial, and lingering response activation from the prime is added to the total accumulated evidence, producing positive priming[1].

Timing and Masking Effects in the NCE Literature

With longer delays between prime and mask in the NCE paradigm, more habituation accrues, which offsets the positive response priming in the evidence accumulation layer, producing negative priming (i.e., an NCE). Similarly, a longer duration prime produces more habituation, and pushes the pattern of results toward a stronger NCE. Finally, if the mask is created through a combination of both directions (i.e., a "relevant" mask), the novel aspects of the mask are more salient (i.e., the line segments that point in the opposite direction to the prime), creating response priming for the incorrect direction. This moves the data pattern even more strongly in the direction of negative priming, producing an NCE regardless of timing. In this manner, the model is similar to the object updating account of Lleras and Enns (2004), although the model can produce a NCE even without a relevant mask. This object updating is the core idea underlying the neural habituation theory—after viewing the prime arrow long enough to produce habituation for the direction of the prime, it is easier to perceive a subsequent stimulus that differs from the prime (i.e., the aspects of the mask that differ from the prime are perceptually salient). Thus, the observation that the NCE is larger with a relevant mask highlights the beneficial effect of habituation (i.e., better perception for the novel aspects of the mask).

In summary, whether arrow priming as measured with reaction times produces a NCE or PCE is determined by the amount of perceptual habituation (which causes NCE) versus response activation (which causes PCE). As prime duration and/or prime-mask ISI increases, the amount of perceptual habituation overtakes the lingering response activation, causing a shift from PCE to NCE when using irrelevant masks or no masks. When using relevant masks, there is response priming for the alternative direction as compared with the prime, due to salient perception of the novel aspects of the mask, and this shifts the pattern to NCE regardless of the prime-mask ISI.

Response Modality and Prime-Target Similarity Effects in the NCE Literature

Several studies in the NCE literature have examined the extent to which primes produce NCE as a function of the similarity between prime and target and whether NCE priming effects are cross-modal. Although we do not formally model these, we outline how the model architecture could be changed to accommodate the task demands of each study, and how this change in architecture would naturally produce the observed results.

The first NCE study to examine prime-target similarity (Eimer & Schlaghecken, 1998) found no NCE when targets consisted of "LL" and "RR," and primes consisted of arrows. In this experiment, subjects were always given letters for targets, and were instructed to respond to "LL" with their left hand, and "RR" with their right hand. The habituation theory could model these results by including both letter perception and direction perception; however, because subjects were never asked to respond to the arrows, only letter perception would be connected to the response layer. As a result, even though perception of the prime and masks would occur as usual, there would be no response priming effects at all, with the response layer only reflecting letter perception.

The LL/RR study can be contrasted with one conducted by Eimer (1999), which used lateral "+" signs as targets presented to the left or right of fixation, requiring a button press for the corresponding hand. Unlike the LL/RR study, this study found a significant NCE for these lateralized + sign targets after viewing arrow primes. A key difference between this study and the LL/RR study is that subjects were given arrow targets on some trials in the + sign study, with the type of target (+ or arrow) occurring randomly across trials, with no warning as to which would occur at the start of the trial. This could be modeled by connecting both arrow perception and + sign position perception to the response layer. Thus, a mask consisting of overlapping arrows would provide response priming for the response direction that differed from the prime direction. This explanation can be applied to other NCE studies using dissimilar primes and targets, such as Experiment 2 of Klapp and Hinkley (2002), which used up-down arrows as targets for some trials, but high-low-pitched tones as targets for other trials (with primes always up-down arrows). They found a cross-modal NCE effect for arrows followed by tones, which is to be expected if the response layer of the habituation model is connected to both arrow perception and tone perception.

Effector modality has been studied by Eimer et al. (2002), who utilized two different types of targets: typical central arrow targets, and lateral symbol targets, similar to the + sign study described above. Primes were always central arrows, and subjects were instructed to respond with different effectors for each target type of target (e.g., feet for lateral symbols but hands for arrows). A

---

[1] For best-fitting parameters, this lingering activation in the response layer was subthreshold, meaning that the response node corresponding to the prime was not actively firing at the time when the target was presented, but the value for the membrane potential was above zero (i.e., the response did not have as far to go to reach threshold), resulting in a larger overall peak magnitude for the target response once the target appeared.

significant NCE was found when primes and targets were both central arrows, but no NCE was found when targets were lateral symbols. This result makes sense considering that there were in fact four possible responses (left-hand, right-hand, left-foot, and right-foot), rather than two. In modeling this result, the habituation model would require four response nodes, with arrow perception mapped to two of the response nodes and symbol position perception mapped to the other two response nodes. Thus, arrow primes would only affect the corresponding response nodes (producing an NCE) but not the other response nodes.

In summary, task demands will dictate the mapping between perception and response. Considering this mapping, priming effects are predicted if the prime and target are connected to the same response nodes. Thus, the NCE reflects an interaction between perceptual and response processes, in light of task demands.

Relationship to Other NCE Theories

Because the habituation model as applied to the NCE paradigm includes both perceptual priming and response priming, it shares aspects with theories on both sides of the debate regarding the question of whether the NCE is a response effect or a perceptual effect. Here we consider these similarities and differences with other accounts of the NCE.

Similar to the object updating account (Lleras & Enns, 2004), the habituation model supposes that the NCE is partly driven by interactions between the features of the prime and the mask (when present). Indeed, the similarity between prime and mask are key to capturing the temporal dynamics of the results, as irrelevant masks generate a different results profile than relevant masks (see Figure 10). The neural habituation model updates its perceptual representations at every simulated millisecond in a manner consistent with the object updating account, but it does not require a stimulus between prime and target (the mask) to produce an NCE. Thus, the habituation model can explain NCE even with irrelevant masks, while also explaining PCE with short duration primes followed by irrelevant masks or no mask. Unlike the object updating account, the habituation model assumes that these positive effects are based in response priming.

In contrast to object updating, the self-inhibition account (Eimer & Schlaghecken, 2002, 2003; Klapp & Hinkley, 2002; Schlaghecken & Eimer, 2002) assumes that the NCE is primarily a response effect. According to this account, an NCE occurs if the prime generates enough activity to trigger self-inhibitory circuits while remaining subliminal. Like the self-inhibition account, an NCE under the neural habituation model requires that the prime generates enough activation to cause habituation (due to synaptic depression). Also, like the self-inhibition model's subliminal requirement, the positive effect of response priming needs to be sufficiently weak according to the habituation model, lest it overcome deficits from habituation. However, the two accounts are reversed in another sense, with the negative component based in response processes for the self-inhibition account but perceptual processes for the habituation model while, at the same time, the potentially offsetting positive component is based in perceptual processes for the self-inhibition account (prime awareness) but response processing (response priming) for the habituation account. That said, because everything is channeled through the response layer in the habituation model, the two accounts are phenomenologically similar.

One way to potentially untangle perception versus response is to examine ERP components that are known to reflect one kind of process or the other. As predicted by the self-inhibition account, arrow priming NCE studies reported lateralized readiness potentials (LRPs) in response to the prime and mask presentations (Eimer & Schlaghecken, 1998; Liu et al., 2014; Praamstra & Seiss, 2005), and the LRP is known to reflect response preparation (Coles, 1989). However, the existence of an LRP does not necessarily indicate that it is the cause of behavior—instead, the root cause of the behavioral deficit might be perceptual processes, with the LRP passively reflecting the effect of these perceptual processes on the response system. Additionally, Verleger et al. (2004), who replicated the LRP results mentioned above, suggested that the polarity reversal that had previously been attributed to self-inhibition (Eimer & Schlaghecken, 1998) could instead reflect activation of the opposite motor cortex caused by the relevant mask. It is therefore entirely possible that, as predicted by the neural habituation model, the perceptual interactions between prime and relevant mask generate positive response priming in the nonprimed direction, which is reflected as activation of the opposite motor cortex and a reversal of LRP polarity.

The self-inhibition account has been formalized with a computational model (Bowman et al., 2006) that shares a few basic aspects with the neural habituation model, such as perceptual dynamics that serve to drive response dynamics. However, the root cause of the NCE in the Bowman model is a special purpose self-inhibition circuit within the response system, similar to motor control in the basal ganglia, rather than a perceptual deficit that weakens response activation. Because this self-inhibition is triggered by the mask, it is not clear how the Bowman model can produce an NCE in the absence of a mask. More generally, whereas the neural habituation model is a wide-ranging theory of perceptual dynamics (it also explains a wide variety of other tasks and stimuli with rapid visual presentations), the self-inhibition model is a specific model of response inhibition in the NCE.

An account of note that shares similarities with both object-updating and self-inhibition is the mask-triggered inhibition account, which proposes that the role of the mask is stopping response accumulation by the prime (Jaskowski, 2008; Jaskowski & Verleger, 2008; Panis & Schmidt, 2016; Schmidt et al., 2015). Like self-inhibition, this account assumes that inhibitory processes targeting the prime-triggered response are responsible for the NCE, and like object-updating, it also assumes that the response-relevant perceptual features of the mask play a significant role in determining the strength of the NCE. A key difference between this account and the self-inhibition account is that it does not require the prime to be effectively masked (Panis & Schmidt, 2016), in fact, mask-triggered inhibition can account for an NCE induced by flankers instead of a mask (Jaskowski, 2008). However, this account cannot explain the presence of an NCE in the absence of a stimulus between prime and target, and, furthermore, there is no formal/mathematical implementation of mask-triggered inhibition.

Like the habituation model, the evaluation window account of Klauer and Dittrich (2010) implicates a higher-level perceptual representation. Also, like the habituation model, deficits in the evaluation window model are relative, rather than reflecting an active inhibition process. More specifically, the evaluation window model supposes that performance reflects the amount of perceptual activation to the target as compared to an initial baseline level. In the case of an evaluation window that excludes the prime, such as occurs with an intervening mask, priming of the wrong answer reduces the baseline starting level for the target, resulting in relatively better performance (i.e., perception of the target is salient given this lowered baseline level). In contrast, an inclusive window uses a baseline level prior to the prime, in which case the activation from the prime is added to that of the target, similar to the response priming explanation of the PCE in the habituation model. However, the evaluation window cannot differentiate between masks types (Klauer & Dittrich, 2010) and, like most NCE accounts (Eimer & Schlaghecken, 1998; Jaskowski, 2008; Lleras & Enns, 2004), it does not provide a quantitative explanation of the gradual transition from PCE to NCE with changes in timing. Instead, the evaluation window model would suppose an abrupt shift from PCE to NCE depending on whether the evaluation window did or did not include the prime. Klauer and Dittrich (2010) addressed this limitation by suggesting that the NCE could be explained by multiple mechanisms, with the evaluation window being one of them; similarly, other authors (Atas et al., 2015) have also suggested that both object updating and self-inhibition could underlie the NCE. In a way, the neural habituation model is as a cohesive implementation of these different mechanisms, directly showing how the varied NCE results can be explained by the combination of perceptual and response priming.

Relationship to Other Accounts of Multistimulus Perception

The overarching goal of the neural habituation model is to bridge across diverse paradigms to provide a unified explanation of cognitive aftereffects. The key aspect of the model is the suppression of repeated stimuli (caused by synaptic depression) as presentation duration or ISI becomes longer. However, the neural habituation model is not the only theory that generalizes across tasks and stimuli; here, we discuss two theories that are similar to the neural habituation model in this regard.

Predictive coding models (Huang & Rao, 2011) propose that the brain learns statistical regularities present in the environment, and uses them to discount predictable components of perceptual input in order to focus on those that are not predictable. In other words, unpredicted features are more salient. Similarly, the habituation model assumes that previously viewed features are weakened due to synaptic depression, which in turn causes novel features to be more salient. Considering a representative framework of predictive coding (Rao & Ballard, 1999), the model is trained to learn statistical regularities from a set of natural images, which it can then apply to subsequent images in order to determine what (in a spatial sense) is predictable and what is salient within a particular image. Analogously, the neural habituation model does something similar on a shorter timescale, determining how previous stimuli in a sequence of events can affect the saliency of subsequent stimuli. With this in mind, the neural habituation model is more similar to temporal implementations of predictive coding (Hosoya et al., 2005), in which predictions are based on dynamic adjustments across time. When the predictive coding model is exposed to a particular stimulus, the model becomes adapted to its features, subtracting them from subsequent stimuli via inhibition; something similar happens in the neural habituation model, only the subtracting is a relative deficit due to synaptic depression. The key difference is that predictive coding is unable to generate positive priming: if something is presented to the model, the only result is discounting (i.e., a negative effect), whereas the neural habituation model predicts a rise-then-fall pattern resulting from the initial increase in activation followed by a slower decrease in activation due to the gradual depletion of neurotransmitter (though in the case of the NCE, positive priming occurred in the neural habituation model because of evidence accumulation in the response layer).

The sometimes opponent process (SOP) model (Wagner, 1981) has primarily been applied to associative learning and conditioning, but the course of a representation activation in this model is remarkably similar to output under neural habituation. According to SOP (Brandon & Wagner, 2002), when a representation is activated by a stimulus presentation, it enters a state of focal processing, which soon decays to a state of weaker secondary activation; once the inducing stimulus is no longer present, the representation decays to inactivity. This pattern of temporal activation is very similar to the one predicted by neural habituation: activation is initially very strong, but soon decreases as a result of synaptic depression; eventually, resource depletion and recovery reach a balance, and activation remains steady (albeit weak) until the stimulus is no longer present, at which point it begins to fade to zero. The key difference between the two accounts is that SOP does not clearly establish the temporal dynamics of the transition between states; instead, transition is determined only by probability parameters (Mazur & Wagner, 1982; Wagner, 1981). This makes SOP unsuitable for predicting the effect of timing on prime effects without the burden of excessive free parameters.

Conclusions

The theory guiding this work assumes that perceptual habituation is useful for temporally segmenting the stream of visual inputs—by habituating to the features of recently viewed objects, new objects are made perceptually salient. This is a result of inhibitory competition influencing the model's perceptual layers: if the representation of recently viewed objects is weaker due to habituation, novel objects will be subjected to less lateral inhibition, and thus will be easier to perceive. In fact, causal evidence that neural habituation enhances novelty detection was recently reported by Jacob and Huber (2020). In their same/different task with word stimuli, they observed a transition from benefits to deficits with increasing duration of the cue word. They collected EEG responses during this task, and a classifier analysis of the trial data indicated that the magnitude of the N400 was highly predictive of "different" responses. Furthermore, the neural habituation model provided a coherent explanation of the perceptual ERP waveforms (e.g., P100 and N170), the N400 response, and the behavioral data, producing repetition deficits following longer duration cue words. Here we ascertained whether these same neural dynamics might underlie repetition deficits in the NCE arrow direction paradigm.

First, we established the perceptual dynamics of orientation perception with a novel orientation priming task that used threshold accuracy as the key dependent measure. Our results revealed a rapid transition from benefits to deficits with increasing duration of an oriented prime stimulus and a second study replicated this effect, while ruling out an explanation in terms of response priming. Unlike the tilt aftereffect, this repetition deficit for orientation perception occurred at a relatively high level considering that the prime and target appeared in different screen locations and were of different spatial frequencies. The neural habituation model was fit to these results, specifying key temporal parameters for orientation perception. Then, with these perceptual aspects of the model fixed, the neural habituation model was augmented with a response layer that accumulated response information throughout the trial sequence, explaining decisional aspects of the task as revealed with reaction times. This augmented model successfully explained the major finding in the NCE literature, reconciling a long-standing debate regard the role of masks in this paradigm. In terms of the NCE, we conclude that NCE deficits reflect perceptual habituation, while benefits (e.g., PCE in the absence of a mask) reflect response priming. More generally, we conclude that the NCE is another example of immediate repetition deficits that arise from neural habituation.

Context

The present research highlights how general properties of neural behavior—for example, synaptic depression—can provide unified accounts of different behavioral paradigms. The neural habituation model's success in explaining the NCE literature suggests that the NCE is a cognitive aftereffect due to neural habituation for higher level forms of perception. This explanation connects the NCE literature to perceptual dynamics and RSVP paradigms more generally, demonstrating links between the NCE and higher level repetition deficits in orthographic, semantic, and face priming (Huber, Tian, et al., 2008; Jacob & Huber, 2020; Potter et al., 2018; Rieth & Huber, 2010, 2017), spatial cuing (Rieth & Huber, 2013), episodic recognition (Huber, Clark, et al., 2008), the attentional blink (Rusconi & Huber, 2018), evaluative priming (Irwin et al., 2010), and semantic satiation (Tian & Huber, 2010, 2013). Similar to these cognitive aftereffects that were previously explained by neural habituation, the NCE in the arrow priming task is explained by neural habituation of orientation perception, which, in turn, reduces activation in the response system. In contrast, positive effects in the arrowing priming task, such as occurs in the absence of an intervening mask, or with an irrelevant intervening mask, are explained by response priming.

References

Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*(5297), 220–224. https://doi.org/10.1126/science.275.5297.221

Atas, A., San Anton, E., & Cleeremans, A. (2015). The reversal of perceptual and motor compatibility effects differs qualitatively between metacontrast and random-line masks. *Psychological Research*, *79*(5), 813–828. https://doi.org/10.1007/s00426-014-0611-3

Bowman, H., Schlaghecken, F., & Eimer, M. (2006). A neural network model of inhibitory processes in subliminal priming. *Visual Cognition*, *13*(4), 401–480. https://doi.org/10.1080/13506280444000823

Brandon, S., & Wagner, A. (2002). Sometimes opponent process (SOP) model of conditioning. In J. H. Byrne, H. Eichenbaum, H. Roediger, & R. F. Thompson (Eds.), *Learning & memory* (2nd ed., pp. 624–628). Macmillan.

Carandini, M., & Heeger, D. (1994). Summation and division by neurons in primate visual cortex. *Science*, *264*(5163), 1333–1336. https://doi.org/10.1126/science.8191289

Chance, F. S., Nelson, S. B., & Abbott, L. F. (1998). Synaptic depression and the temporal response characteristics of V1 cells. *The Journal of Neuroscience*, *18*(12), 4785–4799. https://doi.org/10.1523/jneurosci.18-12-04785.1998

Coles, M. G. H. (1989). Modern mind-brain reading: Psychophysiology, physiology, and cognition *Psychophysiology*, *26*(3), 251–269. https://doi.org/10.1111/j.1469-8986.1989.tb01916.x

D'Ostilio, K., Cremers, J., Delvaux, V., Sadzot, B., & Garraux, G. (2013). Impaired automatic and unconscious motor processes in Parkinson's disease. *Scientific Reports*, *3*(1), 2095. https://doi.org/10.1038/srep02095

Davelaar, E. J., Tian, X., Weidemann, C. T., & Huber, D. E. (2011). A habituation account of change detection in same/different judgments. *Cognitive, Affective and Behavioral Neuroscience*, *11*(4), 608–626. https://doi.org/10.3758/s13415-011-0056-8

Eimer, M. (1999). Facilitatory and inhibitory effects of masked prime stimuli on motor activation and behavioural performance. *Acta Psychologica*, *101*(2–3), 293–313. https://doi.org/10.1016/s0001-6918(99)00009-8

Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1737–1747. https://doi.org/10.1037/0096-1523.24.6.1737

Eimer, M., & Schlaghecken, F. (2002). Links between conscious awareness and response inhibition: Evidence from masked priming. *Psychonomic Bulletin & Review*, *9*(3), 514–520. https://doi.org/10.3758/BF03196307

Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology*, *64*(1-2), 7–26. https://doi.org/10.1016/S0301-0511(03)00100-5

Eimer, M., Schubö, A., & Schlaghecken, F. (2002). Locus of inhibition in the masked priming of response alternatives. *Journal of Motor Behavior*, *34*(1), 3–10. https://doi.org/10.1080/00222890209601926

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743. https://doi.org/10.1038/nn.3689

Fox, E. (1995). Negative priming from ignored distractors in visual selection: A review. *Psychonomic Bulletin & Review*, *2*(2), 145–173. https://doi.org/10.3758/BF03210958

Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, *104*(3), 572–594. https://doi.org/10.1037/0033-295X.104.3.572

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology*, *20*(5), 453–467. https://doi.org/10.1037/h0059826

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *The Cambridge handbook of computational psychology* (pp. 59–100). Cambridge University Press. https://doi.org/10.1017/CBO9780511816772.006

Hosoya, T., Baccus, S. A., & Meister, M. (2005). Dynamic predictive coding by the retina. *Nature*, *436*(7047), 71–77. https://doi.org/10.1038/nature03689

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Cognitive Science*, *2*(5), 580–593. https://doi.org/10.1002/wcs.142

Huber, D. E., Clark, T. F., Curran, T., & Winkielman, P. (2008). Effects of repetition priming on recognition memory: Testing a perceptual fluency-disfluency model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1305–1324. https://doi.org/10.1037/a0013370

Huber, D. E., & O'Reilly, R. C. (2003). Persistence and accommodation in short-term priming and other perceptual paradigms: Temporal segregation through synaptic depression. *Cognitive Science*, *27*(3), 403–430. https://doi.org/10.1016/S0364-0213(03)00012-0

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Quach, R. (2002). Mechanisms of source confusion and discounting in short-term priming 2: Effects of prime similarity and target duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1120–1136. https://doi.org/10.1037/0278-7393.28.6.1120

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Ruys, K. I. (2001). Perception and preference in short-term word priming. *Psychological Review*, *108*(1), 149–182. https://doi.org/10.1037/0033-295X.108.1.149

Huber, D. E., Shiffrin, R. M., Quach, R., & Lyle, K. B. (2002). Mechanisms of source confusion and discounting in short-term priming: 1. Effects of prime duration and prime recognition. *Memory & Cognition*, *30*(5), 745–757. https://doi.org/10.3758/bf03196430

Huber, D. E., Tian, X., Curran, T., O'Reilly, R. C., & Woroch, B. (2008). The dynamics of integration and separation: ERP, MEG, and neural network studies of immediate repetition effects. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(6), 1389–1416. https://doi.org/10.1037/a0013625

Irwin, K. R., Huber, D. E., & Winkielman, P. (2010). Automatic affective dynamics: An activation-habituation model of affective assimilation and contrast. In T. Nishida, L. C. Jain, & C. Faucher (Eds.), *Modeling machine emotions for realizing intelligence. Smart innovation, systems and technologies* (pp. 17–34). Springer, Berlin. https://doi.org/10.1007/978-3-642-12604-8_2

Jacob, L. P. L., & Huber, D. E. (2020). Neural habituation enhances novelty detection: An EEG study of rapidly presented words. *Computational Brain & Behavior*, *3*(2), 208–227. https://doi.org/10.1007/s42113-019-00071-w

Jaskowski, P. (2008). The negative compatibility effect with nonmasking flankers: A case for mask-triggered inhibition hypothesis. *Consciousness and Cognition*, *17*(3), 765–777. https://doi.org/10.1016/j.concog.2007.12.002

Jaskowski, P., & Przekoracka-Krawczyk, A. (2005). On the role of mask structure in subliminal priming. *Acta Neurobiologiae Experimentalis*, *65*(4), 409–417.

Jaskowski, P., & Verleger, R. (2008). What determines the direction of subliminal priming. *Advances in Cognitive Psychology*, *3*(1-2), 181–192. https://doi.org/10.2478/v10053-008-0024-1

Jin, D. Z., Dragoi, V., Sur, M., & Seung, H. S. (2005). Tilt aftereffect and adaptation-induced changes in orientation tuning in visual cortex. *Journal of Neurophysiology*, *94*(6), 4038–4050. https://doi.org/10.1152/jn.00571.2004

Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, *27*(2), 117–143. https://doi.org/10.1016/0010-0277(87)90016-3

Klapp, S. T. (2005). Two versions of the negative compatibility effect: Comment on Lleras and Enns (2004). *Journal of Experimental Psychology: General*, *134*(3), 431–435. https://doi.org/10.1037/0096-3445.134.3.431

Klapp, S. T., & Hinkley, L. B. (2002). The negative compatibility effect: Unconscious inhibition influences reaction time and response selection. *Journal of Experimental Psychology: General*, *131*(2), 255–269. https://doi.org/10.1037/0096-3445.131.2.255

Klauer, K. C., & Dittrich, K. (2010). From sunshine to double arrows: An evaluation window account of negative compatibility effects. *Journal of Experimental Psychology: General*, *139*(3), 490–519. https://doi.org/10.1037/a0019746

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*(3), 856–867. https://doi.org/10.1152/jn.1994.71.3.856

Lakens, D., & Caldwell, A. (2019). *Simulation-based power-analysis for factorial ANOVA designs*. https://osf.io/pn8mc/

Liu, P., Chen, X., Dai, D., Wang, Y., & Wang, Y. (2014). A subliminal inhibitory mechanism for the negative compatibility effect: A continuous versus threshold mechanism. *Experimental Brain Research*, *232*(7), 2305–2315. https://doi.org/10.1007/s00221-014-3925-x

Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *Journal of Experimental Psychology: General*, *133*(4), 475–493. https://doi.org/10.1037/0096-3445.133.4.475

Lleras, A., & Enns, J. T. (2005). Updating a cautionary tale of masked priming: Reply to Klapp (2005). *Journal of Experimental Psychology: General*, *134*(3), 436–440. https://doi.org/10.1037/0096-3445.134.3.436

Lleras, A., & Enns, J. T. (2006). How much like a target can a mask be? Geometric, spatial, and temporal similarity in priming: A reply to Schlaghecken and Eimer (2006). *Journal of Experimental Psychology: General*, *135*(3), 495–500. https://doi.org/10.1037/0096-3445.135.3.495

Marr, D. C., & Poggio, T. (1977). *From understanding computation to understanding neural circuitry*. https://dspace.mit.edu/handle/1721.1/5782

Mazur, J. E., & Wagner, A. R. (1982). An episodic model of associative learning: Acquisition. In M. Commons, R. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Acquisition* (pp. 3–39). Ballinger.

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1155–1172. https://doi.org/10.1037/0278-7393.18.6.1155

Panis, S., & Schmidt, T. (2016). What is shaping RT and accuracy distributions? Active and selective response inhibition causes the negative compatibility effect. *Journal of Cognitive Neuroscience*, *28*(11), 1651–1671. https://doi.org/10.1162/jocn_a_00998

Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., & Chelazzi, L. (2018). Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. *Journal of Vision*, *18*(10), 1049. https://doi.org/10.1167/18.10.1049

Pelli, D. G., & Bex, P. (2013). Measuring contrast sensitivity. *Vision Research*, *90*, 10–14. https://doi.org/10.1016/j.visres.2013.04.015

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, *31*(7–8), 1337–1350. https://doi.org/10.1016/0042-6989(91)90055-a

Potter, K. W., Donkin, C., & Huber, D. E. (2018). The elimination of positive priming with increasing prime duration reflects a transition from perceptual fluency to disfluency rather than bias against primed words. *Cognitive Psychology*, *101*, 1–28. https://doi.org/10.1016/j.cogpsych.2017.11.004

Praamstra, P., & Seiss, E. (2005). The neurophysiology of response competition: Motor cortex activation and inhibition following subliminal response priming. *Journal of Cognitive Neuroscience*, *17*(3), 483–493. https://doi.org/10.1162/0898929053279513

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Rawji, V., Modi, S., Latorre, A., Rocchi, L., Hockey, L., Bhatia, K., Joyce, E., Rothwell, J. C., & Jahanshahi, M. (2020). Impaired automatic but intact volitional inhibition in primary tic disorders. *Brain*, *143*(3), 906–919. https://doi.org/10.1093/brain/awaa024

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, *95*(3), 318–339. https://doi.org/10.1037/0033-295X.95.3.318

Rieth, C. A., & Huber, D. E. (2010). Priming and habituation for faces: Individual differences and inversion effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 596–618. https://doi.org/10.1037/a0018737

Rieth, C. A., & Huber, D. E. (2013). Implicit learning of spatiotemporal contingencies in spatial cueing. *Journal of Experimental Psychology:*

Human Perception and Performance, *39*(4), 1165–1180. https://doi.org/10.1037/a0030870

Rieth, C. A., & Huber, D. E. (2017). Comparing different kinds of words and word-word relations to test an habituation model of priming. *Cognitive Psychology*, *95*, 79–104. https://doi.org/10.1016/j.cogpsych.2017.04.002

RStudio Team. (2020). RStudio: Integrated development for R. http://www.rstudio.com/

Rusconi, P., & Huber, D. E. (2018). The perceptual wink model of non-switching attentional blink tasks. *Psychonomic Bulletin & Review*, *25*(5), 1717–1739. https://doi.org/10.3758/s13423-017-1385-6

Schlaghecken, F., & Eimer, M. (2002). Motor activation with and without inhibition: Evidence for a threshold mechanism in motor control. *Perception & Psychophysics*, *64*(1), 148–162. https://doi.org/10.3758/BF03194564

Schlaghecken, F., Klapp, S. T., & Maylor, E. A. (2009). Either or neither, but not both: Locating the effects of masked primes. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1656), 515–521. https://doi.org/10.1098/rspb.2008.0933

Schlaghecken, F., Rowley, L., Sembi, S., Simmons, R., & Whitcomb, D. (2008). The negative compatibility effect: A case for self-inhibition. *Advances in Cognitive Psychology*, *3*(1–2), 227–240. https://doi.org/10.2478/v10053-008-0027-y

Schmidt, T., Hauch, V., & Schmidt, F. (2015). Mask-triggered thrust reversal in the negative compatibility effect. *Attention, Perception & Psychophysics*, *77*(7), 2377–2398. https://doi.org/10.3758/s13414-015-0923-4

Sekuler, R., & Littlejohn, J. (1974). Tilt aftereffect following very brief exposures. *Vision Research*, *14*(1), 151–152. https://doi.org/10.1016/0042-6989(74)90133-3

Singer, J. H., & Diamond, J. S. (2006). Vesicle depletion and synaptic depression at a mammalian ribbon synapse. *Journal of Neurophysiology*, *95*(5), 3191–3198. https://doi.org/10.1152/jn.01309.2005

Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision Research*, *44*(12), 1297–1320. https://doi.org/10.1016/j.visres.2004.01.002

Stenner, M. P., Baumgaertel, C., Heinze, H. J., Ganos, C., & Müller-Vahl, K. R. (2018). Intact automatic motor inhibition in patients with Tourette syndrome. *Movement Disorders*, *33*(11), 1800–1804. https://doi.org/10.1002/mds.27493

Sumner, P., Nachev, P., Morris, P., Peters, A. M., Jackson, S. R., Kennard, C., & Husain, M. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron*, *54*(5), 697–711. https://doi.org/10.1016/j.neuron.2007.05.016

The MathWorks, Inc. (2015). MATLAB (Version 2015a) [Computer software]. https://www.mathworks.com/

Tian, X., & Huber, D. E. (2010). Testing an associative account of semantic satiation. *Cognitive Psychology*, *60*(4), 267–290. https://doi.org/10.1016/j.cogpsych.2010.01.003

Tian, X., & Huber, D. E. (2013). Playing “duck duck goose” with neurons: Change detection through connectivity reduction. *Psychological Science*, *24*(6), 819–827. https://doi.org/10.1177/0956797612459765

Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(2), 719–723. https://doi.org/10.1073/pnas.94.2.719

Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *The Journal of Neuroscience*, *24*(46), 10440–10453. https://doi.org/10.1523/JNEUROSCI.1905-04.2004

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592. https://doi.org/10.1037/0033-295X.108.3.550

Usher, M., Olami, Z., & McClelland, J. L. (2002). Hick’s law in a stochastic race model with speed-accuracy tradeoff. *Journal of Mathematical Psychology*, *46*(6), 704–715. https://doi.org/10.1006/jmps.2002.1420

Verleger, R., Jas¨kowski, P., Aydemir, A., Van Der Lubbe, R. H. J., & Groen, M. (2004). Qualitative differences between conscious and nonconscious processing? On inverse priming induced by masked arrows. *Journal of Experimental Psychology: General*, *133*(4), 494–515. https://doi.org/10.1037/0096-3445.133.4.494

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Erlbaum.

Weidemann, C. T., Huber, D. E., & Shiffrin, R. M. (2005). Confusion and compensation in visual perception: Effects of spatiotemporal proximity and selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 40–61. https://doi.org/10.1037/0096-1523.31.1.40

Weidemann, C. T., Huber, D. E., & Shiffrin, R. M. (2008). Prime diagnosticity in short-term repetition priming: Is primed evidence discounted, even when it reliably indicates the correct answer? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 257–281. https://doi.org/10.1037/0278-7393.34.2.257

Zucker, R. S., & Regehr, W. G. (2002). Short-term synaptic plasticity. *Annual Review of Physiology*, *64*(1), 355–405. https://doi.org/10.1146/annurev.physiol.64.092501.114547

Received July 16, 2020
Revision received January 8, 2021
Accepted January 16, 2021 ■