

From Sunshine to Double Arrows: An Evaluation Window Account of Negative Compatibility Effects

Karl Christoph Klauer and Kerstin Dittrich
Albert-Ludwigs-Universität Freiburg

In category priming, target stimuli are to be sorted into 2 categories. Prime stimuli preceding targets typically facilitate processing of targets when primes and targets are members of the same category, relative to the case in which both stem from different categories, a positive compatibility effect (PCE). But negative compatibility effects (NCEs) are also sometimes observed. An evaluation window account (Klauer, Teige-Mocigemba, & Spruyt, 2009) of PCE and NCE in evaluative priming (category good versus category bad) is applied to masked arrow priming (Eimer & Schlaghecken, 1998; category left versus category right). Key principles of the account are that participants evaluate incoming evidence across a time window, and decisions about stimulus category are driven by changes in evidence weighted according to the Weber-Fechner law, leading to NCE for primes falling outside the time window and PCE for primes inside the time window. In Experiments 1–4, factors considered obligatory for NCE by current accounts of arrow priming are successively removed; yet, NCE remained intact as predicted by the evaluation window account. Furthermore, the evaluation window account, but none of the current accounts, predicts NCE without a stimulus intervening between prime and target at intermediate prime–target stimulus-onset asynchrony (Experiment 5) and when target onset comes as a surprise (Experiment 6). We conclude that the evaluation window account describes a hitherto overlooked mechanism that contributes to PCE and NCE in arrow priming and that it appears to generalize beyond the confines of evaluative priming to the diverse class of category-priming paradigms.

Keywords: negative compatibility effects, arrow priming, evaluative priming, category priming, response control

In recent years, there has been immense interest in adapting response-time paradigms from cognitive psychology to the purpose of measuring attitudes unobtrusively. One of the most prominent of the new methods is the evaluative priming procedure (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Evaluative priming is a classical instance of category priming. In each trial of the paradigm, two stimuli, a prime and a target, are presented in close succession. Targets carry strong evaluations (e.g., the words sunshine and death might appear as targets on different trials), and the participants' task is to categorize targets as denoting something good or bad as fast and as accurately as possible. A typical finding is that prime category (i.e., whether the prime denotes something good or bad) interacts with the speed and accuracy with which targets are sorted into the categories good and bad. Specifically, a positive compatibility effect (PCE) is typically observed (Klauer & Musch, 2003), so that responses to targets are facilitated if prime and target share the same category membership, relative to the case in which prime and target stem from different categories. When the

respondent's evaluation of the prime as good or bad is unknown, this finding can be exploited to infer the prime evaluation based on the effect that the prime has on target processing via the assumption of PCE (Wittenbrink, 2007).

One problem with this inference is that negative compatibility effects (NCE) have occasionally been reported in evaluative priming as reviewed by Klauer et al. (2009). Thus, responses to targets were sometimes delayed when prime and target share the same category membership, relative to when prime and target stem from different categories. Because the validity of evaluative priming as a measurement tool hinges on the assumption that primes engender PCE rather than NCE, it was important in this research tradition to characterize the conditions under which PCE was to be expected and the separate conditions leading to NCE. Klauer et al. (2009) proposed and tested an evaluation window account¹ of NCE in the evaluative priming paradigm that achieved just this.

The assumptions and mechanisms invoked in that account are not confined to the evaluative-priming paradigm. In principle, the evaluation window account applies to any kind of priming paradigm in which response tendencies elicited by prime and target can be compatible or incompatible. This opens a wide range of relevant compatibility paradigms and compatibility effects studied for a variety of applied and theoretical reasons in diverse fields of inquiry within psychology, in social psychology (De Houwer,

Karl Christoph Klauer and Kerstin Dittrich, Institut für Psychologie, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany.

The research reported in this article was supported by Deutsche Forschungsgemeinschaft Grant Kl 614/31-1 to Karl Christoph Klauer. Both authors contributed equally to this article.

Correspondence concerning this article should be addressed to Karl Christoph Klauer, Institut für Psychologie, Albert-Ludwigs-Universität Freiburg, D-79085 Freiburg, Germany. E-mail: christoph.klauer@psychologie.uni-freiburg.de

¹ Klauer et al. (2009) referred to the evaluation window account as the psychophysical account.

2003), in psycholinguistics (e.g., Forster, Mohan, & Hector, 2003), in research on subliminal perception and processing (e.g., Greenwald, Draine, & Abrams, 1996), in motivation and emotion research (e.g., Rothermund, 2003), in studying the mechanisms of cognitive control of speeded reactions (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001), in developmental psychology (e.g., Degner & Wentura, 2010), and in many other fields.

The purpose of the present article is to investigate whether the evaluation window account successfully predicts phenomena beyond the narrow confines of evaluative priming. For that purpose, we turn to what is perhaps the *Drosophila melanogaster* of research on NCE, the arrow-priming paradigm² (Eimer & Schlaghecken, 1998), and test whether the mechanism described in the evaluation window account contributes to shaping NCE and PCE in arrow priming. Research on arrow priming has generated a rich empirical and theoretical literature (for reviews see Jaśkowski & Verleger, 2007, and Sumner, 2007, among others). Moreover, arrow priming is perhaps as dissimilar from evaluative priming as possible within the range of category-priming paradigms as elaborated below. For these reasons, success in the present enterprise would constitute a first step in raising one's confidence in the possibility that the evaluation window account may indeed possess a scope of explanation as large as suggested in the previous paragraph.

Arrow Priming

In 1998, Eimer and Schlaghecken discovered a NCE in a masked priming paradigm that instigated a considerable amount of empirical research and theoretical debate (e.g., Jaśkowski & Verleger, 2007; Sumner, 2007). In Eimer and Schlaghecken's (1998) Experiment 1A, participants saw double arrowheads pointing left or right. Participants were to indicate the left or the right direction. Prior to the target stimulus, two additional stimuli were presented in quick succession, a prime and a mask. Like the target, the prime was a double arrowhead (there were also neutral primes), whereas the mask was formed by the two target shapes overlaid on one another. As confirmed by a forced-choice test of prime visibility, the mask rendered prime visibility low.

Responses to the target were faster and more accurate when prime and target arrows pointed in different directions than when they pointed in the same direction. In other words, NCE was found. In contrast, the typical finding in priming paradigms of this kind is PCE, and this is in particular true for masked priming (e.g., Dehaene et al., 1998; Forster et al., 2003; Klauer, Eder, Greenwald, & Abrams, 2007; Klauer, Musch, & Eder, 2005; Reynvoet, Gevers, & Caessens, 2005, among others). NCEs are therefore not only counterintuitive but also exceptional. That the effect occurs even though the prime is masked suggests also that the effect may reveal something about the unconscious and spontaneous regulation of motoric responses.

Subsequent research established the robustness of the effect and identified many factors moderating it (e.g., Jaśkowski & Verleger, 2007). For example, the effect hinges on prime–target stimulus-onset asynchrony (SOA) being sufficiently long; the effect is more pronounced for so-called relevant masks that contain features of possible targets than for so-called irrelevant masks without such feature overlap. NCE does, however, also occur for irrelevant masks of different kinds (Klapp, 2005; Schlaghecken & Eimer,

2006; Sumner, 2007; Sumner & Husain, 2008). As pointed out by Sumner (2007), different masks typically also differ in how effectively they mask the prime, and NCE tends to decrease as prime visibility increases, as elaborated in the General Discussion. Furthermore, the effect is stronger when all three stimuli (prime, mask, and target) are presented centrally than when the stimuli are presented at different locations.

On the theoretical side, a number of innovative theories have been proposed to account for NCE in arrow priming, perhaps the most prominent being the accounts in terms of self-inhibition (Eimer & Schlaghecken, 1998; Schlaghecken & Eimer, 2000, 2002, 2006), mask-triggered inhibition (Jaśkowski, 2007; Jaśkowski & Przekoracka-Krawczyk, 2005), and object updating (Lleras & Enns, 2004, 2005, 2006), as elaborated below.

Despite numerous differences between these theories, there is general agreement on two points: (a) Priming effects reflect the regulation and inhibitory control of motor activation elicited by primes, masks, and targets (Boy & Sumner, in press). That is, the effect is believed to reflect the interaction of motor-response tendencies evoked by the prime or the prime-mask ensemble, on the one hand, and the target, on the other hand. (b) In addition, there is general agreement that a stimulus (i.e., the mask) has to intervene between prime and target for NCE to occur. In fact, PCE reliably occurs when there is no mask (e.g., Jaśkowski, 2008; Klapp & Hinkley, 2002; Lleras & Enns, 2004; Verleger, Jaśkowski, Aydemir, van der Lubbe, & Groen, 2004). In an interesting twist on this, Boy, Clarke, and Sumner (2008) compared a condition in which a mask intervened between prime and target with a condition in which the mask was present before the prime, but at a different location, and was then gradually moved to mask the prime at that point in time at which it appeared in the first condition. NCE was observed in the first condition; PCE was observed in the second condition.

Given that NCE in masked arrow priming is perhaps the most well-known and best investigated instance of NCE, the question asked in this article is whether Klauer et al.'s (2009) evaluation window account also plays a role in shaping NCE in arrow priming. However, the pattern of findings obtained in arrow priming makes it clear that more than the evaluation window account must be involved, as elaborated below, and thus, we do not claim that the evaluation window account is viable as a stand-alone account of arrow priming. Instead, our argument will be that factors such as described in current theories contribute to shaping priming effects in arrow priming but that there is also a contribution best described by the evaluation window account that has been overlooked so far.

The Evaluation Window Account

The evaluation window account (Klauer et al., 2009) acknowledges that regulation and control of motor activation elicited by primes, masks, and targets contributes to prime–target compatibility effects (Klauer et al., 2005; Reynvoet et al., 2005). However, it

² Whereas Eimer and Schlaghecken (1998) used arrows pointing left or right as stimuli, subsequent research used a variety of stimuli, in most cases, pictorial stimuli with spatial (e.g., left versus right or up versus down) features in addition to arrows, but we refer to this research as research on arrow priming for lack of a better term.

postulates that compatibility effects additionally include a component that is divorced from both early perceptual analysis of the stimuli and motor control stages. The component operates at a more abstract, semantic level of categorizing stimuli according to the task-relevant categories, that is, as pointing left, versus pointing right, in the arrow-priming paradigm³ and as good, versus bad, in the evaluative-priming paradigm. Thus, we assume that compatibility effects are caused both at the motor level by interactions of motor activations elicited by primes and targets and at the categorization level in terms of interactions of bits of information about the appropriate categorization elicited by primes and targets. It is not necessary for the present purposes to map out a detailed architecture for the relations between processes at both levels, but one possibility is that as information about stimulus category accrues, it is simultaneously fed into the motor system and translated into motor activation for whatever motor response is mapped on the stimulus category by the task instructions (i.e., processes at both levels operate as cascading processes; McClelland, 1979). Another possibility is that in the presence of strong sensorimotor links, early perceptual analysis of the stimulus may be sufficient to engender some amount of motor activation directly (Kunde, Kiesel, & Hoffmann, 2003), independently of, and in addition to, parallel deeper processing of stimulus category.

For the categorization level, the key principles of the evaluation window account are (a) that participants evaluate incoming evidence about stimulus category across a time window, termed the evaluation window, irrespective of source (prime, mask, or target) of the evidence, (b) that decisions about stimulus category are driven by changes in evidence weighted according to the Weber-Fechner law (e.g., Miller, 1964), such that small changes are more apparent against a background of little evidence to date than against a background of much evidence, and (c) that participants use recent experience and context to synchronize the evaluation window with the stream of incoming stimulus events.

Specifically, with regard to the first principle (a), it is assumed that separate counters are set up for the different task-relevant categories (e.g., for the categories left and right in arrow priming). These counters accumulate incoming evidence for the respective category; that is, they provide an ongoing tally of the activation of category-relevant features as provided by perceptual analysis of incoming information. It is important to note that they do so irrespective of the source of that evidence (e.g., irrespective of whether the evidence was elicited by a prime or the target). In the absence of incoming information, counter states decay and gradually return to a zero state of no activation.

In speeded decision tasks, participants are assumed to base their categorization on detecting increases in the counter states over a specific time period, the evaluation interval. Increases monitored over a brief time period following target onset afford a correct and fast categorization for unambiguous stimuli.⁴ Using increases over a short evaluation interval as the basis for one's decision is less error prone than using the counter states at some fixed point in time shortly after target onset, given that the counter states at any given time may still vary strongly as a function of activation fed into them by irrelevant sources prior to target onset.

With regard to the second principle (b), we assume that it is difficult to synchronize the evaluation window with the stream of ongoing stimulus events. Because the counters do not record the source of incoming activation, the onset of the evaluation window

has to be synchronized closely with the target onset to exclude activation from irrelevant sources that feed activation into the counters prior to target onset. Yet, under time pressure, there is pressure to open the evaluation window as soon as possible to arrive at a fast decision. It is assumed that participants solve this problem imperfectly and that they tend to open the evaluation window anticipatorily somewhat prior to target onset. Participants thereby ensure that they can make immediate use of the target-derived evidence as soon as any is available. However, the evaluation window will typically thereby include some of the activation driven by primes preceding targets with short SOA, leading to PCE. Evaluation windows in which prime-derived activation contributes to changes in counter states within the window will be referred to as inclusive evaluation windows.

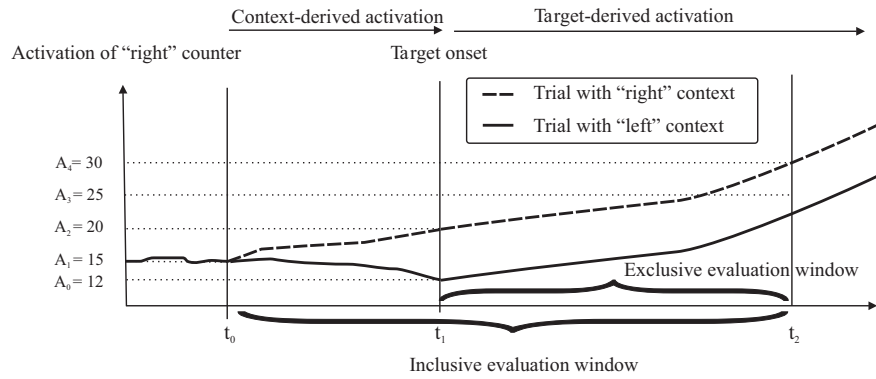
Inspired by the Weber-Fechner law of psychophysics, it is assumed that detecting an increase in the evaluation window is easier, that is, faster and more accurate, starting out from a low level of counter activation at window onset than from a high level. This leads to NCE for primes that are effectively excluded from the evaluation window, that is, for primes that have contributed to increasing the counter state up to window onset but have not contributed to counter increases in the evaluation window itself: If an excluded prime shares the target's category, the counter for that category is already at a high level at window onset, making it difficult to detect an additional target-driven increase in the evaluation window. In contrast, if an excluded prime is not a member of the target category, the counter for the target's category will start out from a comparatively lower level at window onset, making it easier to identify the target's category. The result is NCE for primes excluded from the evaluation window. Evaluation windows in which prime-derived activation does not contribute to changes in the counter states within the window are referred to as exclusive evaluation windows.

Thus, for an inclusive evaluation window that starts at prime onset or shortly thereafter, PCE is expected. As the window onset moves away from prime onset toward target onset, PCE diminishes, crosses the zero point, and eventually turns into NCE for exclusive evaluation windows that exclude most of the prime-derived evidence but include the target-derived evidence. Figure 1 illustrates the prediction of PCE and NCE for an inclusive evaluation window and an exclusive evaluation window, respectively.

With regard to the third principle (c), it is proposed that participants capitalize on repetitive and predictive sequences of stimulus events to synchronize window onset. In particular, when trials are structured the same, presenting, say, a prime, an intervening stimulus, and a target in close succession with fixed temporal intervals, the last stimulus before the target reliably predicts the impending

³ Because there are only two target stimuli in the original arrow priming paradigm, categorization can be thought of as identification in this case; that is, the categories have a size of one.

⁴ We assume that the two counters are negatively correlated or mutually inhibitory so that activation can accumulate quickly in the appropriate counter for unambiguous stimuli that support only one categorization. But both for neutral stimuli and for ambiguous stimuli (such as masks composed of two superimposed arrows pointing in opposite directions), little change occurs in either counter.



Evaluation window	“Right” context	“Left” context	Context effect
Inclusive (from t_0 to t_2):	$\frac{\Delta A}{A_1} = \frac{30-15}{15} = 1.00$	$\frac{\Delta A}{A_1} = \frac{22-15}{15} = 0.46$	0.53 (PCE)
Exclusive (from t_1 to t_2):	$\frac{\Delta A}{A_2} = \frac{30-20}{20} = 0.50$	$\frac{\Delta A}{A_0} = \frac{22-12}{12} = 0.83$	-0.33 (NCE)

Figure 1. Activation of the counter for the category right given a right-pointing target for a trial with right context (e.g., a right-pointing prime) and a trial with left context (e.g., a left-pointing prime). The context stimulus (e.g., a prime) sets on at t_0 , followed by the target at t_1 . Two evaluation windows are shown; an inclusive window begins at t_0 , an exclusive window at t_1 . The trajectory of activation for the trial with right context is represented by the broken line, the trajectory for the trial with left context by the continuous line. The right context leads to a counter increase from A_1 to A_2 from context onset (t_0) to target onset (t_1), whereas the left context leads to a decrease from A_1 to A_0 , due to lateral inhibition (see Footnote 4). Participants base their decision on the increase ΔA in activation that accrues during the evaluation window relative to the initial counter state A . PCE = positive compatibility effect; NCE = negative compatibility effect; t = time.

onset of the target. It is assumed that the last stimulus before the target will therefore come to be used as a go signal for opening the evaluation window. This leads to an exclusive evaluation window for the prime and to an inclusive evaluation window for the intervening stimulus. Accordingly, NCE should be observed for the prime.

Empirical support for these different assumptions has so far been gathered in the evaluative-priming paradigm in social-cognitive research (Klauer et al., 2009). For example, in a paradigm with two primes followed by a target (e.g., the three words *sunshine*, *death*, and *joy* might be presented in close succession), NCE was found for the effects of the first prime on the target, and PCE was found for the effects of the second prime on the target. This agrees well with the idea sketched in the last paragraph that the first prime will be excluded from the evaluation window and that the second prime will be included. In addition, NCE was observed even when no stimulus intervened between prime and target in conditions promoting exclusive evaluation windows in different ways, as elaborated below.

Applied to arrow priming, participants learn to use the mask in partitioning the stream of incoming stimuli and come to use it as a go signal for opening the evaluation window. Hence, the prime is likely to fall out of the evaluation window. The evaluation window account can thereby account for the basic NCE phenomenon in arrow priming.

Current Accounts of Arrow Priming

In this section, we briefly review the major current accounts of NCE in arrow priming: the account by self-inhibition, the account by mask-triggered suppression, and the account by object updating.

The Account by Self-Inhibition

According to the account by self-inhibition, NCE reflects inhibition of motor activation elicited by the prime (Eimer & Schlaghecken, 1998; Schlaghecken & Eimer, 2000, 2002, 2006). Inhibition is thought to occur automatically after an initial phase of motor activation (a) if perceptual evidence for the prime is quickly removed, usually by a subsequent mask, and (b) if the strength of the sensory representation of the prime is nevertheless sufficiently large. Inhibition causes NCE. Note in particular that in line with the first condition (a), PCE is obtained when there is no stimulus between prime and target (e.g., Jaśkowski, 2008; Klapp & Hinkley, 2002; Lleras & Enns, 2004; Verleger et al., 2004).

The Account by Mask-Triggered Inhibition

In the account by mask-triggered inhibition (Jaśkowski, 2007; Jaśkowski & Przekoracka-Krawczyk, 2005), it is assumed that inhibition is triggered by the mask, more precisely, by any stim-

ulus immediately following the prime and appearing in the focus of attention, provided it does not support the perceptual hypothesis concerning the prime's identity. The ongoing response activation will thereby be inhibited, leading to NCE. Mask-triggered inhibition is assumed to be particularly strong for masks that contain features of possible targets (Jaśkowski, 2008). Thus, NCE should be smaller for masks that do not contain features of possible targets (e.g., when the mask consists of horizontal and vertical lines) than for masks that do (e.g., when the mask is composed of superimposed left-pointing and right-pointing arrowheads). Masks of the latter kind are often called relevant; masks of the former kind are often called irrelevant. In fact, NCE is more pronounced for relevant masks than for irrelevant masks (e.g., Lleras & Enns, 2004, 2005, 2006; Schlaghecken & Eimer, 2006).

The Account by Object Updating

In the account by object updating (Lleras & Enns, 2004, 2005, 2006), NCE is thought to reflect perceptual interactions of prime and mask. When prime and mask follow each other in quick succession, they will be perceived as one changing stimulus, and the aspects that change will be especially salient. In particular, for the relevant mask with superimposed left-pointing and right-pointing arrowheads, what changes in going from prime to mask is that arrowheads pointing in the direction opposite from that of the prime are added. If elements are found that call for another response, as in this case, the prime-triggered activation stops and activation of the alternative response begins. It is this latter response tendency that is responsible for NCE. NCE is thus really a PCE relative to the updated information and the response tendency associated with it. Conversely, little NCE should occur for irrelevant masks for which the updated information is not associated with any of the task responses.

The account by object updating can thereby explain the effect of mask relevance described in the previous section, but it does not account for NCE with irrelevant masks (see the above section, Arrow Priming). Lleras and Enns (2005, 2006) elaborated on the account by object updating by adding two additional ideas: One is onset-triggered suppression leading to some amount of response inhibition whenever a nontarget is presented; the second is a repeated-location advantage. The former principle accounts for NCE occurring for irrelevant masks, the latter accounts for the fact that NCE tends to be larger for primes, masks, and targets presented at the same location than for primes, masks, and targets presented at different locations (e.g., Lleras & Enns, 2006; Schlaghecken & Eimer, 2000). With these additions, the account by object updating overlaps considerably with the account by mask-triggered inhibition.

The Evaluation Window Account in Relation to Current Accounts

Applied to arrow priming, the mechanism described by the evaluation window account differs markedly from the current accounts. In the evaluation window account, one component of NCE is located at a more central and abstract level of categorizing the target into the task-relevant categories (such as the categories of arrows pointing left, versus right) than at the level of interacting motor-response tendencies. In addition, according to the evalua-

tion window account, it should be possible to obtain NCE even when no stimulus (such as a mask) intervenes between prime and target under certain conditions. The role of the mask is not to remove the perceptual evidence provided by the prime (as in the account by self-inhibition), not to trigger inhibition of motor activation (as in the account by mask-triggered inhibition), and not to interact perceptually with the prime (as in the account by object updating). Instead, it serves as a kind of go signal for an early and anticipatory positioning of the onset of the evaluation window that thereby excludes the prime.

The evaluation window account can thereby account for NCE in arrow priming. It is, however, not viable as a stand-alone account. For example, the account does not readily explain moderating effects of mask relevance: NCE should be observed largely irrespective of which particular mask is shown, as long as it can be used effectively to partition the stream of incoming stimuli, and there is little reason to believe that irrelevant and relevant masks should differ pronouncedly in their usefulness as go signals for opening the evaluation window.

We acknowledge that perceptual interactions of prime, mask, and factors related to motor control play a major role in shaping PCE and NCE in arrow priming. Stimuli are presented in quick succession, making it likely that they are sometimes perceived as one changing stimulus, as postulated in the account by object updating. In addition, simple response mappings, often mapping two stimuli on two responses in a spatially congruent manner (e.g., a left double arrow is mapped on a left key, and a right double arrow is mapped on a right key), are used making it likely that participants have stored or quickly acquired strong and direct sensorimotor links, calling for some machinery capable of controlling and down regulating potentially misleading motor activation triggered by primes. In addition, there are only a few primes and targets, sometimes only one per task category, minimizing the role of the categorization stage and, thereby, the contribution described by the evaluation window account.

In consequence, it is not reasonable to claim that the evaluation window account provides an alternative account replacing these previous accounts. Instead, we propose that the evaluation window account describes yet another mechanism, hitherto overlooked, that additionally shapes PCE and NCE in arrow priming. Note also that the purpose of the present research was not to discriminate between the current accounts of NCE in arrow priming, although we discuss implications of our findings for these accounts when appropriate.

Our research strategy is twofold: In a first series of experiments (Experiments 1 to 4), we stick relatively closely to the typical procedures of the arrow-priming paradigm and successively eliminate factors thought to be responsible for NCE in that paradigm according to current theories, factors that are, however, not obligatory according to the evaluation window account. We show that a residual amount of NCE remains that is difficult to account for by the current accounts but that is well accounted for by the evaluation window account.

In Experiments 5 and 6, we depart more boldly from the traditional arrow-priming paradigm, in that we present primes and targets without intervening stimuli. None of the existing theories of arrow priming is capable of explaining NCE without a mask or without at least a nonmasking third stimulus intervening between prime and target. Nevertheless, the evaluation window account

demarcates conditions under which NCE should nevertheless be observed. Experiments 5 and 6 implemented two different sets of such conditions.

Experiment 1

In the first experiment, we aimed to replicate NCE in arrow priming. We thus presented three stimuli per trial that we refer to as prime, Stimulus 2, and target, respectively. Primes and targets were double arrowheads pointing either left or right. According to the evaluation window account, it should make little difference whether Stimulus 2 effectively masks the prime. Consequently, we used presentation parameters for which the prime was clearly visible (see also Jaśkowski, 2008). Like in previous studies, Stimulus 2 was sometimes a relevant mask (superimposed arrows) and sometimes an irrelevant mask. A new feature was that we also included trials in which prime and Stimulus 2 themselves were double arrowheads pointing left or right. This is inspired by experiments in which two primes were shown in succession prior to the target in the evaluative-priming paradigm (Fockenberg, Koole, & Semin, 2008; Gawronski, Deutsch, & Seidel, 2005; Klauer et al., 2009), and NCE was observed for the first prime.

According to the evaluation window account, participants learn to use Stimulus 2 in partitioning the stream of incoming stimuli into separate perceptual episodes and come to use it as a go signal for setting up the evaluation window. Hence, the prime is likely to fall outside of the evaluation window. The stand-alone prediction of the evaluation window account is that NCE should occur for trials with relevant masks, with irrelevant masks, and with response-valent Stimuli 2 (i.e., with left-pointing or right-pointing double arrowheads as Stimulus 2) and that NCE should be of similar size in all of these conditions.

Method

Participants. Participants were 20 University of Freiburg students with different majors; mean age was 24 years, ranging from 19 years to 42 years. Participants received a monetary compensation for participating that was contingent on their performance. For each correct response that occurred within a given response window in one of the experimental blocks, they received 1 Euro cent (equivalent to U.S.\$1.23). They could thereby earn a maximum of 7.68 Euros (equivalent to U.S. \$9.42).

Materials. Primes and targets were double arrowheads pointing left or right. Two additional symbols could occur as Stimulus 2. The relevant mask consisted of the left and right arrow stimuli superimposed; the irrelevant mask consisted only of horizontal and vertical lines and conveyed no directional information. The stimuli were similar to those shown in Figure 2 (see row for Experiment 2), but the arrow stimuli in Experiment 1 had an aperture of 45° rather than 90°, as in Figure 2. All stimuli subtended an area of approximately 1.2° × 1.2° in terms of visual angle. They were presented in black on a gray background in the center of a 48.3 cm CRT screen with 100 Hz refresh rate.

Procedure. In each trial, prime, Stimulus 2, and target were presented in quick succession in the center of the screen. Prime and Stimulus 2 were each presented for 20 ms, followed by an empty interval of 40 ms, resulting in an SOA of 60 ms between prime and Stimulus 2 and between Stimulus 2 and target.

There were two practice blocks of 20 trials each and 16 experimental blocks of 48 trials each. In the experimental blocks and in the second practice block, the target remained on screen until 600 ms had passed or a response was entered, whichever event occurred earlier. In the first practice block, the target remained on screen until a response was given. The intertrial interval was 500 ms.









Prime / Target	Stimulus 2	Experiment
		2
		3, 4
		5, 6
		

Figure 2. Stimuli used in Experiments 2–6.

The participants' task was to categorize the target as left or right as quickly and as accurately as possible by pressing one of two response keys. Participants used the interior keys of two computer mice positioned left and right in front of them (Voss, Leonhart, & Stahl, 2007).

In the first practice block, participants received trialwise feedback of their response latency as well as error feedback (following a false response the word *Fehler* [error] was shown). Both bits of feedback were presented at screen center for 1,000 ms. In the subsequent blocks, trialwise feedback was omitted, and participants received end-of-block feedback of mean response latency, percentage correct responses, and amount of money earned (only experimental blocks) in the last block. Participants were tested in individual sessions of about 45 min.

Design. The factors prime (left arrow versus right arrow), Stimulus 2 (left arrow, right arrow, relevant mask, irrelevant mask), and target (left arrow versus right arrow) were crossed orthogonally, resulting in 16 different trials. Each of these was shown three times per experimental block. In the practice blocks, factor combinations were randomly sampled from the 16 possible combinations.

Results

The accuracy data (proportion correct scores) were arcsine-transformed for the statistical analyses, but we retransform results back to the original percentage scale in reporting mean values in the body of the text and in the tables in the Appendix for ease of interpretation in this and the subsequent experiments. NCEs in the accuracy domain are shown on the arcsine scale in the figures that allows us to plot the appropriate confidence intervals. Degrees-of-freedom in F tests were Greenhouse-Geisser adjusted, when applicable, in all analyses in this article.

Response latencies below 150 ms were omitted, as were outliers in each individual's distribution of latencies as identified by Tukey's criterion (i.e., latencies that were below the first quartile minus 1.5 times the interquartile range or above the third quartile plus 1.5 times the interquartile range; Clark-Carter, 2004, Chapter 9). This led to the exclusion of 2.21% of the trials. Mean latency was 385 ms; mean error rate was 2.79%.

For the sake of readability, we present analyses in terms of compatibility effects (signed so that positive values indicate PCE and negative values indicate NCE). Analyses of correct response latencies and percentage-correct scores are presented in the Appendix.

Figure 3 shows the compatibility effects in the latency domain (upper panel) and in the accuracy domain (lower panel). The response-valent Stimuli 2 were coded as compatible with the target (i.e., both Stimulus 2 and target are the same double arrow), versus incompatible (i.e., Stimulus 2 and target are different double arrows). As can be seen, there was strong NCE in trials with relevant masks, smaller NCE in trials with irrelevant masks, and PCE in trials with response-valent compatible Stimuli 2.

The prime-target compatibility effects were submitted to analyses of variance with factor Stimulus 2 (irrelevant, relevant, incompatible, and compatible). This revealed a strong effect of Stimulus 2 both in the latency domain, $F(3, 57) = 100.76, p < .01, \epsilon = .65, \eta_p^2 = .84$, and in the accuracy domain, $F(3, 57) = 22.81, p < .01, \epsilon = .89, \eta_p^2 = .55$.

Follow-up t tests revealed that NCE was significantly larger (a) in trials with relevant mask than in trials with irrelevant mask, and

(b) in trials with irrelevant mask than in trials with response-valent Stimuli 2 (collapsing over the two response-valent Stimuli 2), both in the latency and the accuracy domain, all $t(19) > 3.04, p < .01$. Across trials with response-valent Stimuli 2, PCE was observed in the latency domain and in the accuracy domain ($M = 17$ ms and $M = 2.81\%$, respectively), both $t(19) > 14.50, p < .01$.

An interesting pattern emerged for response-valent Stimuli 2: In the latency domain, there was PCE in trials with compatible response-valent Stimuli 2 and NCE in trials with incompatible Stimuli 2. Both prime-target compatibility effects were individually significant (see Figure 3) as was the difference between them, $t(19) = 14.50, p < .01$. In contrast, in the accuracy domain, there was PCE for both kinds of response-valent Stimulus 2, with PCE individually significant for incompatible Stimulus 2 (see Figure 3) and significantly larger for incompatible Stimulus 2 than for compatible Stimulus 2, $t(19) = 2.52, p = .02$.

Discussion

In Experiment 1, NCE was observed for trials with relevant and irrelevant masks, and PCE was observed for trials with response-valent Stimuli 2. Given that the prime was clearly visible and perceptual evidence for it was not immediately removed, the observed NCE is probably difficult to explain in terms of the account by self-inhibition (see also Jaśkowski, 2007, 2008). Note that we did not implement a formal test of prime visibility, a fact that subtracts from the conclusiveness of our claim here. But prime visibility is beyond any reasonable doubt for the presentational parameters used in the subsequent experiments.

According to the evaluation window account, NCE should have been observed largely irrespective of which particular Stimulus 2 is shown, as long as it can be used effectively to partition the stream of incoming stimuli. With hindsight, this use of Stimulus 2 may be difficult when Stimulus 2 itself is a double arrowhead and identical to either the prime stimulus or the target stimulus: The fast, successive presentation of two identical stimuli may be perceived as one flickering stimulus rather than as two distinct perceptual episodes. But both relevant and irrelevant masks should be effective in partitioning the stream of incoming stimuli. The evaluation window account therefore cannot explain the strong effect of mask relevance: NCE was substantially stronger for the relevant mask than for the irrelevant mask. We attribute the effects of mask relevance to perceptual interactions between prime and Stimulus 2 as per object updating and/or to stronger recruitment of control processes by the relevant mask than by the irrelevant mask, as per mask-triggered inhibition.

The pattern observed for response-valent Stimuli 2 is broadly consistent with the simple idea that identical stimuli, repeated in close succession, are perceived as one flickering stimulus. For example, according to the account by object updating and the account by mask-triggered suppression, little updating and suppression, respectively, would occur for the almost immediate repetition of identical stimuli, so that response activation can proceed more or less unhindered.

In trials with response-valent compatible Stimulus 2 (identical Stimulus 2 and target), this leads to a head start in correct response activation of 120 ms for compatible primes (repetition of three identical stimuli; head start equals the prime-Stimulus 2 SOA, 60 ms, plus the target-Stimulus 2 SOA, 60 ms), but of only 60 ms for

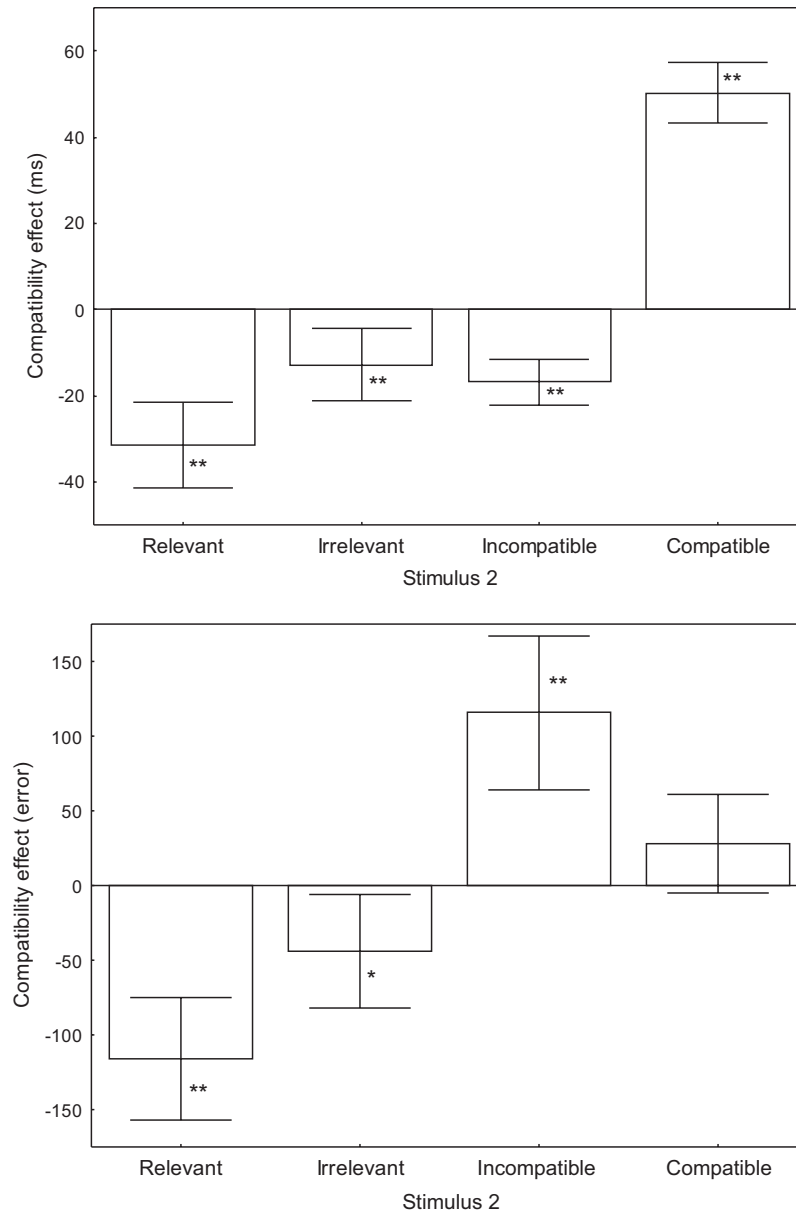


Figure 3. Experiment 1. Compatibility effects in the latency domain (upper panel) and in the accuracy domain (lower panel, based on arcsine-transformed values multiplied by 1,000) as a function of Stimulus 2 (relevant, irrelevant, incompatible, compatible). Error bars show the 95% confidence intervals. * $p < .05$. ** $p < .01$.

incompatible primes (only Stimulus 2, but not the prime, already activates the correct response), a difference of 60 ms. Mean latency was 300 ms for the first condition and was 351 ms for the second, a difference of 51 ms. Similarly, mean latency for trials with response-valent compatible Stimulus 2 was 326 ms, and mean latency for trials with response-valent incompatible Stimulus 2 (no head start) was 432 ms, a difference of 110 ms, much of which is accounted for by the mean head start of 90 ms in trials with compatible Stimulus 2. Simultaneously, the repetition of identical Stimulus 2 and target in trials with compatible Stimuli 2 raised accuracy almost to ceiling ($M = 99.87\%$ and $M = 87.19\%$, for compatible Stimuli 2 and incompatible Stimuli 2, respectively), so

that the prime had little chance to engender a pronounced compatibility effect in the accuracy domain in trials with compatible Stimuli 2, accounting for the smaller PCE in trials with compatible Stimuli 2 relative to incompatible Stimuli 2.⁵

⁵ This line of reasoning does not explain the individually significant NCE observed in trials with incompatible Stimulus 2 in the latency domain. Note, however, (a) that none of the accounts considered can explain NCE in the latency domain accompanied by PCE in the accuracy domain and (b) that although NCE was also observed in the comparable condition of Experiment 2, it was not individually significant in that experiment.

Experiment 2

Perceptual interactions between the stimuli should be reduced as stimulus durations and blank intervals between stimuli are increased. For one group of participants in Experiment 2, termed the slow group, stimulus durations and blank intervals between stimuli were therefore lengthened by a factor of 3. In the other group, termed the fast group, presentation conditions were as in Experiment 1.

At the more leisurely pace of presentation of the slow group, all stimuli should be clearly perceived as separate stimuli even when they have the same identity. Therefore, the last symbol before the target should consistently be used as go signal for positioning the evaluation window, leading to an exclusive evaluation window for the prime in all trials. This gives rise to Predictions a and b: Inasmuch as perceptual interactions between prime and Stimulus 2, as described by object updating, are reduced when the stimuli are not perceived as one changing scene but are perceived as separate episodes, (a) the effects of mask relevance should be reduced in the slow group. In addition, (b) NCE should then occur even for response-valent Stimuli 2 in the slow group according to the evaluation window account.

Method

Participants were 60 University of Freiburg students with different majors; mean age was 24 years, ranging from 18 years to 42 years. One participant was an extreme outlier, according to Tukey, with a mean reaction time of 514 ms in the total sample's distribution of mean response latencies ($M = 363$ ms, $SD = 37.64$). This participant was excluded from the analysis.

Procedures and design were identical to Experiment 1, except for the following differences. Stimuli are shown in Figure 2. Participants were randomly assigned to two different groups. Conditions for members of the fast group were identical to those of Experiment 1. In the slow group, SOA between and durations of prime and Stimulus 2 were multiplied by 3, resulting in an SOA of 180 ms and in durations of 60 ms for each of prime and Stimulus 2. Thus, any two stimuli were separated by a blank interval of 120 ms in the slow group.

Results

Response latencies were preprocessed as in Experiment 1, excluding 3.45% of trials. Mean latency was 360 ms; mean error rate was 1.47%. In Figure 4, compatibility effects are presented as a function of Stimulus 2 and group in the latency domain (upper panel) and in the accuracy domain (lower panel). As can be seen, NCE prevailed in most conditions, and the effect of Stimulus 2 was visibly dampened in the slow group relative to the fast group.

The compatibility effects were submitted to analyses of variance with factors Stimulus 2 (within-participants) and group (between-participants). This revealed a main effect of group in the latency domain, $F(1, 57) = 6.64$, $p = .01$, $\eta_p^2 = .10$, and in the accuracy domain, $F(1, 57) = 5.11$, $p = .03$, $\eta_p^2 = .08$, reflecting overall larger NCE in the slow group than in the fast group. There was also a main effect of Stimulus 2 in the latency domain, $F(3, 171) = 139.42$, $p < .01$, $\varepsilon = .86$, $\eta_p^2 = .71$, and in the accuracy domain, $F(3, 171) = 16.46$, $p < .01$, $\varepsilon = .86$, $\eta_p^2 = .22$. These main effects

were moderated by a significant interaction of group and Stimulus 2: $F(3, 171) = 51.18$, $p < .01$, $\varepsilon = .86$, $\eta_p^2 = .47$, and $F(3, 171) = 10.21$, $p < .01$, $\varepsilon = .86$, $\eta_p^2 = .15$, in the latency domain and in the accuracy domain, respectively.

Follow-up t tests were conducted to assess Predictions a and b. With regard to the first prediction (a), in the fast group, a significant effect of mask relevance was again observed in the latency domain: NCE was larger in trials with relevant mask than in trials with irrelevant mask, $t(27) = -2.52$, $p = .02$ (see Figure 4 for means). In contrast, in the slow group, the effect of mask relevance was leveled, ($t < 1$). There were no significant effects of mask relevance in the accuracy domain (in both groups, $t < 1$).

With regard to the second prediction (b), in the fast group, PCE was again observed in trials with response-valent Stimuli 2 ($M = 23$ ms and $M = 2.02\%$ in the latency domain and accuracy domain, respectively), both $ts(28) > 3.09$, $p < .01$. In contrast, in the slow group, NCE was observed in these trials, which was significant both in the latency domain ($M = -9$ ms), $t(30) = -2.57$, $p = .02$, and in the accuracy domain ($M = -0.64\%$), $t(30) = -2.65$, $p = .01$. In the slow group, NCE in these trials was significantly smaller than NCE in trials with irrelevant mask in the latency domain, $t(30) = -3.57$, $p < .01$, whereas the difference was not significant in the accuracy domain, $t(30) = -1.32$, $p = .20$.

In the fast group, there was again PCE for response-valent compatible Stimuli 2 and NCE for incompatible Stimuli 2 in the latency domain (see Figure 4), and the difference between the two compatibility effects was again significant, $t(27) = 11.30$, $p < .01$. In the accuracy domain, significant PCE was again found for incompatible Stimuli 2, and the effect of Stimulus 2 compatibility missed significance in a two-tailed t test, $t(27) = -1.91$, $p = .07$. In contrast, in the slow group, NCE was larger for trials with incompatible Stimulus 2 than for trials with compatible Stimulus 2 in both the latency and the accuracy domain, both $ts(30) > 3.21$, $p < .01$ (see Figure 4). Analyses of variance of response latencies and percentage-correct scores as a function of prime, Stimulus 2, target, and group are presented in the Appendix.

Discussion

In Experiment 2, stimuli were presented at a slower pace in a slow group, making it easier to separate them perceptually. As argued above, we believe that perceptual factors (object updating) and factors related to motor control (mask-triggered inhibition) have a role in shaping NCE in arrow priming and that they are in particular responsible for the modulating effect of mask relevance and for the effect of whether Stimulus 2 is itself response-valent.

In contrast, the evaluation window account suggests that the nature of Stimulus 2 is less important as long as it can be used effectively to partition the stream of events into separate perceptual episodes. This led to the expectation that the Stimulus 2 influence should be reduced as perceptual interactions are reduced, that is, in the slow group. In fact, the effect of mask relevance was leveled in that group, and NCE, instead of PCE, was observed for response-valent Stimuli 2.

This is consistent with the evaluation window account and with the assumption that perceptual interactions between prime and Stimulus 2 play a role as postulated by object updating. It may also be possible to explain this pattern of results in terms of mask-triggered inhibition. The account by mask-triggered inhibition

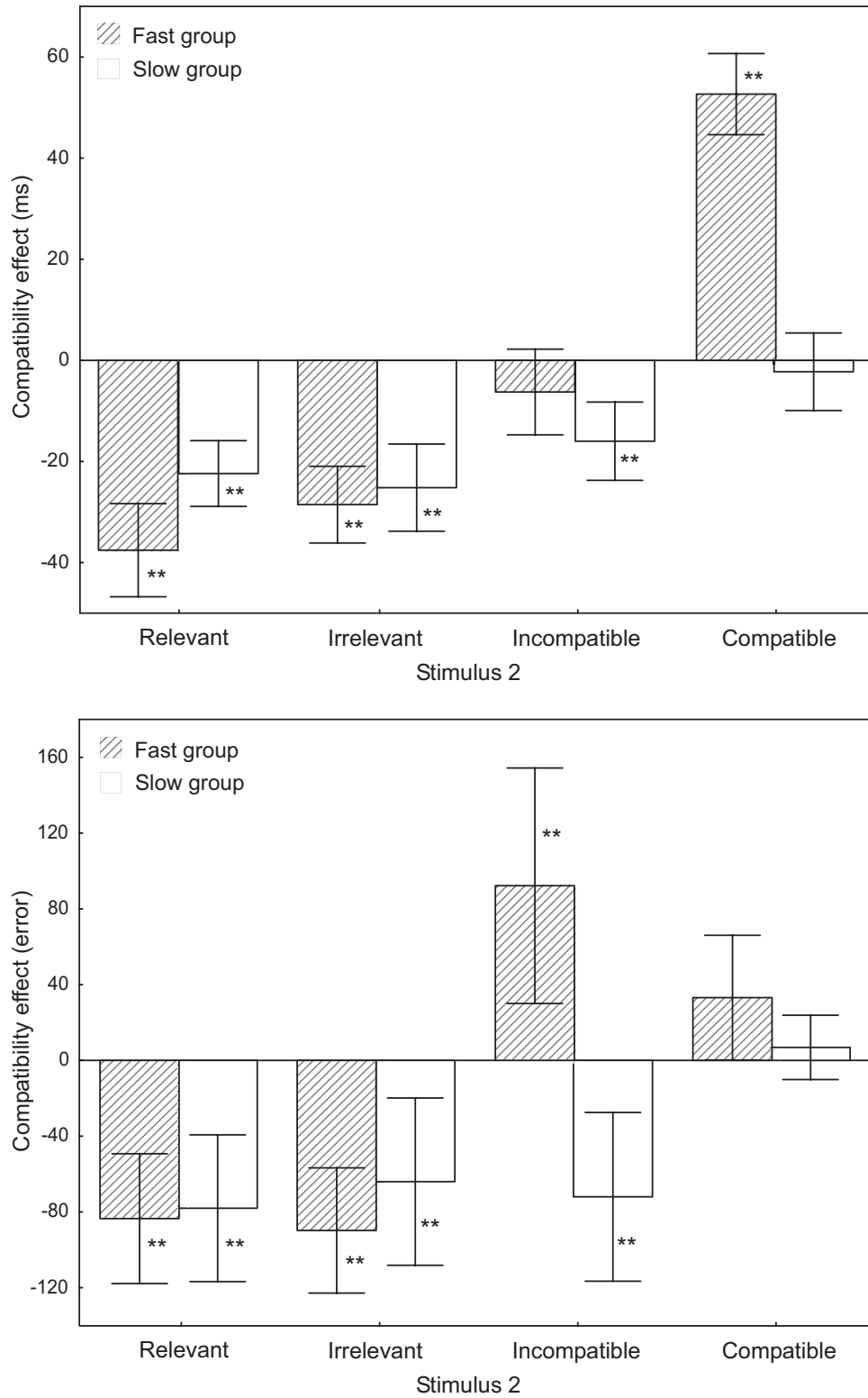


Figure 4. Experiment 2. Compatibility effects in the latency domain (upper panel) and in the accuracy domain (lower panel, based on arcsine-transformed values multiplied by 1,000) as a function of Stimulus 2 (relevant, irrelevant, incompatible, compatible) and group (fast, slow). Error bars show the 95% confidence intervals. * $p < .05$. ** $p < .01$.

postulates two temporal gradients: the response tendency evoked by the prime decay over time and the strength of mask-triggered inhibition is a function of the time elapsed between Stimulus 2 and target. These factors can easily account for overall shifts in the size of NCE as temporal aspects of stimulus presentation are changed, that is, as a function of group in the present case. Depending on the functional shape assumed for the gradients, it is probably possible to account for the leveling of the effect of mask relevance in the slow group and for the transition from PCE to NCE for response-valent Stimuli 2. It is unclear how these accounts would explain the effects of Stimulus 2 compatibility for response-valent Stimuli 2 because the predictions of these accounts have not been spelled out for response-valent Stimuli 2.

Although the effects of Stimulus 2 were dampened in the slow group and the effect of mask relevance, in particular, was eliminated, there remained effects of Stimulus 2: NCE was significantly smaller in trials with response-valent Stimuli 2 than in trials with irrelevant (and relevant) masks; and in trials with response-valent Stimuli 2, NCE was significantly larger for incompatible Stimuli 2 than for compatible ones.

Experiment 3

According to the present line of argument, the remaining Stimulus 2 effects reflect remaining perceptual interactions of prime and Stimulus 2 and/or the impact of factors related to motor control. In Experiment 3, we further reduced the potential for perceptual interactions between prime and Stimulus 2 by using more clearly dissimilar primes and Stimuli 2. In particular, stimuli from two stimulus families were mixed.

One stimulus family, the arrow family, consists of the same stimuli used in Experiment 2. The other stimulus family, the letter family, consists of letter symbols based on the letters *r* and *l*. Targets and primes were double letters *ll* and *rr*, the former to be responded to with the response “left,” the latter to be responded to with the response “right” (see also Eimer & Schlaghecken, 1998, Experiment 1B). The relevant mask consisted of these two stimuli superimposed; the irrelevant mask was a partially filled square of the same size as the relevant mask (see Figure 2).

As before, participants saw sequences of three stimuli: prime, Stimulus 2, and target. In one group, termed the single-family group, all three stimuli stemmed from the same family, either the arrow family or the letter family, with stimulus family chosen randomly for each trial. In the other group, termed the mixed-family group, stimulus families were mixed within trials: Prime and target were always from the same randomly sampled family, whereas Stimulus 2 stemmed from the other family. In this group, Stimulus 2 is thus clearly and consistently discernible as a separate stimulus bearing little resemblance to prime and target. This should support its use in partitioning the sequence of stimuli into different perceptual episodes as per the evaluation window account, leading to consistently exclusive response windows. Furthermore, the impact of perceptual interactions between prime and Stimulus 2 as per object updating should be further reduced given that the stimulus features that change in going from prime to Stimulus 2 are no longer as unambiguously related to the target as before.

Thus, a leveling of the Stimulus 2 effects in the mixed-family group would be consistent with object updating and the evaluation window account. Note, however, that for both groups, the relevant

masks contain features of possible targets, whereas the irrelevant masks do not, and so they continue to qualify as relevant masks and irrelevant masks, respectively, according to the account by mask-triggered suppression.

A number of studies have had symbolic stimuli such as digits, letters, words, and pictures in masked-priming paradigms and usually showed only PCE (e.g., Dehaene et al., 1998; Kiesel, Kunde, Pohl, & Hoffmann, 2006; Klauer et al., 2007; Klauer et al., 2005; Reynvoet et al., 2005; but see Bennett, Lleras, Oriet, & Enns, 2007). Many of these have used what would be classified as irrelevant masks in the present context, along with conditions that impede the usefulness of the mask for perceptually separating the prime from the target (such as sandwich masking in which the prime is embedded in a forward and a backward mask). Kiesel, Berner, and Kunde (2008) implemented the presentational parameters typical of arrow priming relatively closely, using symbolic stimuli (digits) along with arrow stimuli in different experiments. They found NCE for symbolic stimuli with relevant masks but not with irrelevant masks, suggesting a strong effect of mask relevance for symbolic stimuli, although the comparison was one between experiments.

Method

Participants. Participants were 40 University of Freiburg students with different majors; mean age was 23 years, ranging from 20 years to 30 years. One participant was an extreme outlier, according to Tukey, with mean error rate of 17.04% in the total sample’s distribution of error rates ($M = 2.25\%$, $SD = 0.50$). This participant was excluded from the analysis.

Materials. Stimuli used in this experiment are displayed in Figure 2. In addition to left and right double arrows, we used the double letters *ll* and *rr* mapped on the responses “left” and “right,” respectively. As can be seen in Figure 2, the letters were symmetrical along the vertical axis so that they did not contain spatial features biasing a left versus right decision. They were of approximately the same size as the arrow stimuli. There were also a new relevant mask and a new irrelevant mask for the letter family; the relevant mask consisted of both letter symbols superimposed, one on the other. The irrelevant mask was a five-by-five matrix with random white and black cells and the same height and width as the relevant mask (see Figure 2).

Procedure and design. Procedures and design were the same as for the slow group in Experiment 2, except for the following differences. Participants were randomly assigned to two different groups. Members of the single-family group saw stimuli from only one family in each trial; that is, all three stimuli, prime, Stimulus 2, and target, belonged to either the arrow family or the letter family. The stimulus family was randomly chosen for each trial, with half the trials within one block presenting stimuli from the arrow family and half presenting stimuli from the letter family.

The participants of the mixed-family group saw stimuli from both families in each trial. In particular, prime and target belonged to the same stimulus family, whereas the intervening Stimulus 2 belonged to the other stimulus family. Within each block, prime and target were sampled from the arrow family for half the trials; they were sampled from the letter family for the other half. These trials were randomly mixed within each block.

In both groups, a total of 12 experimental blocks with 64 trials each were administered. The factors prime category (left vs. right),

Stimulus 2 category (irrelevant mask, relevant mask, left, and right), target category (left vs. right), and target family (arrow vs. letter) were crossed orthogonally, resulting in 32 combinations of factor levels. In each block, each of the 32 combinations was represented two times. Two additional practice blocks consisted of 48 trials each in this and the following experiments, presenting factor combinations that were randomly sampled from the 32 possible combinations detailed above.

Results

Response latencies were preprocessed as in Experiment 1, excluding 3.00% of the trials. Mean correct response latency was 379 ms; mean error rate was 1.93%.

In Figure 5, compatibility effects are shown for the single-family group (left panels) and the mixed-family group (right panels) as a function of Stimulus 2 category and stimulus family in

the latency domain (upper panels) and in the accuracy domain (lower panels). Consider the single-family group first. For the arrow family, NCE prevailed as in the comparable slow group in Experiment 2, with relatively little effect of mask relevance but with smaller NCE and little sign of compatibility effects for trials with, respectively, response-valent incompatible and compatible Stimuli 2. For the letter family, there was a strong effect of mask relevance and of Stimulus 2 compatibility with PCE in the latency domain for the irrelevant mask and the response-valent compatible Stimuli 2. In contrast, in the mixed-family group, NCE was more uniformly observed across the different kinds of Stimulus 2 and stimulus families.

The compatibility effects were submitted to analyses of variance with factors Stimulus 2 category, stimulus family (both factors within-participants), and group (between-participants). For the latency domain, this revealed a significant interaction of Stimulus 2

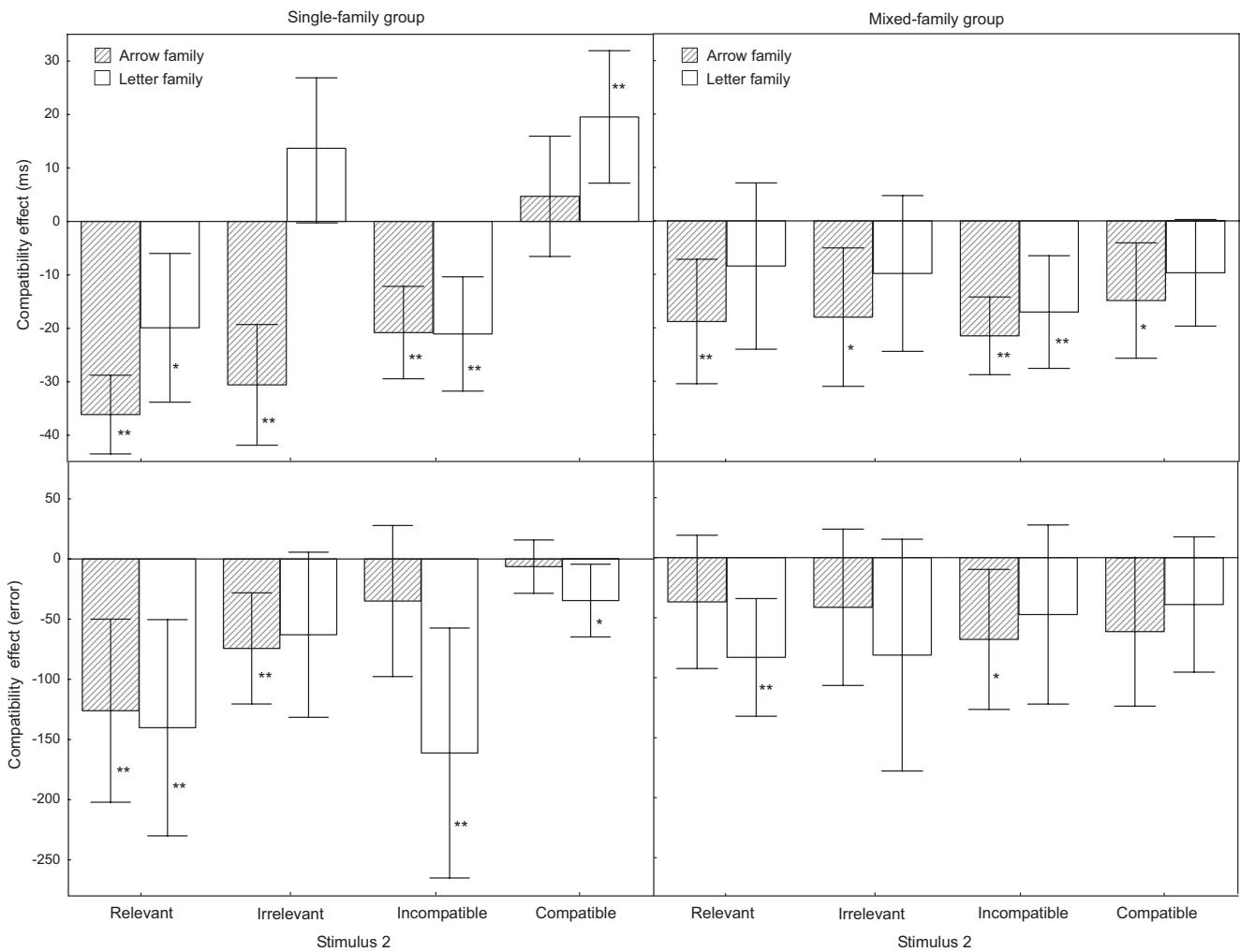


Figure 5. Experiment 3. Compatibility effects in the single-family group (left panels) and the mixed-family group (right panels) in the latency domain (upper panels) and in the accuracy domain (lower panels, based on arcsine-transformed values multiplied by 1,000) as a function of Stimulus 2 (relevant, irrelevant, incompatible, compatible) and stimulus family of prime and target (arrow, letter). Error bars show the 95% confidence intervals. * $p < .05$. ** $p < .01$.

category and group, $F(3, 111) = 9.57, p < .01, \epsilon = .77, \eta_p^2 = .21$; for the accuracy domain, there were no significant effects or interactions (largest $F = 2.70$, smallest $p = .11$). Separate analyses of variance were next computed for the single-family group and the mixed-family group.

Consider the single-family group first. The interaction of Stimulus 2 category and stimulus family was significant in the latency domain, $F(3, 57) = 7.83, p < .01, \epsilon = .81, \eta_p^2 = .29$; for the accuracy domain, $F(3, 57) = 1.71, p = .08, \epsilon = .69, \eta_p^2 = .08$. Follow-up t tests contrasted (a) relevant and irrelevant mask, (b) irrelevant mask and response-valent Stimuli 2, and (c) response-valent compatible with incompatible Stimuli 2, separately for each stimulus family.

For the arrow family, the results from the slow group of Experiment 2 were replicated: There was no effect of mask relevance, $|t| < 1$, and $t(19) = -1.21, p = .24$, in the latency and accuracy domains, respectively (see Figure 5 for means). NCE was, however, again significantly larger in trials with irrelevant mask than in trials with response-valent Stimuli 2; $t(19) = -3.15, p < .01$, and $t(19) = -2.19, p = .04$, for latency and accuracy domain, respectively. Finally, NCE was again significantly larger in trials with response-valent incompatible Stimuli 2 than in trials with compatible Stimuli 2 in the latency domain, $t(19) = -3.89, p < .01$, whereas the difference was not significant in the accuracy domain, $|t| < 1$.

For the letter family, there was a strong effect of mask relevance in the latency domain, $t(19) = -3.29, p < .01$; for the accuracy domain, $t(19) = -1.18, p = .25$. But compatibility effects were not significantly different for irrelevant mask and response-valent Stimuli 2: $t(19) = 1.53, p = .14$, and $t(19) = 0.70, p = .49$, for latency domain and accuracy domain, respectively. In addition, NCE was again significantly larger in trials with response-valent incompatible Stimuli 2 than in trials with compatible Stimuli 2, $t(19) = -6.71, p < .01$, and $t(19) = -2.56, p < .01$, in the latency domain and accuracy domain, respectively. Taken together, the results of Experiment 2 were replicated in the comparable condition (single-family group, arrow-family trials). In particular, there were again remaining effects of Stimulus 2.

Consider next the mixed-family group. In the analyses of variance, there was neither an interaction of stimulus family and Stimulus 2 category nor a main effect of Stimulus 2 category ($F < 1$) for both the latency and the accuracy domain. Only the main effect of stimulus family began to approach significance in the latency domain, $F(1, 18) = 2.68, p = .12, \eta_p^2 = .13$, with descriptively larger NCE when prime and target stemmed from the arrow family than when they stemmed from the letter family ($F < 1$, for the accuracy domain). NCE was significantly different from zero in the latency domain for arrow and letter family ($M = -18$ ms), $t(18) = -4.63, p < .01$, and ($M = -11$ ms), $t(18) = -2.77, p = .01$, respectively. NCE was also individually significant in the accuracy domain for both stimulus families ($M = -0.97\%$), $t(18) = -2.71, p < .01$, and ($M = -3.05\%$), $t(18) = -3.90, p = .01$, respectively. Analyses of variance of response latencies and percentage-correct scores as a function of prime category, Stimulus 2 category, target category, stimulus family, and group are presented in the Appendix.

Discussion

In Experiment 3, we introduced a new stimulus family based on letters instead of arrows. In the single-family group, conditions in arrow trials were comparable with those of the slow group in the previous experiment, and the results pattern was replicated: The effect of mask relevance was leveled, but there remained substantial Stimulus 2 effects on prime-induced NCE: NCE was larger for masks than for response-valent Stimuli 2 and larger for response-valent incompatible Stimuli 2 than for compatible Stimuli 2. The former effect was absent in the letter trials, but there was an effect of mask relevance in these trials. The effect of mask relevance for the symbolic letter stimuli conceptually replicates an analogous finding by Kiesel et al. (2008), as described in the introduction to this experiment within one experiment.

The idea of Experiment 3 was to create a condition, through the use of two stimulus families and mixed-family trials, in which perceptual interactions between primes and targets would be further reduced. The prediction was that any influence of Stimulus 2 type on prime-induced NCE effects should thereby be dampened further. The results confirmed this major prediction: Whereas there were effects of Stimulus 2 in the single-family group as just reviewed, such effects were completely leveled in the mixed-family group.

This pattern of results is consistent with the evaluation window account (augmented by the possibility of perceptual interactions of prime and target) and the account by object updating (augmented by the idea of onset-triggered suppression). Results may be more difficult to explain in terms of mask-triggered inhibition. The different kinds of Stimulus 2 do not differ in terms of whether they contain targetlike features across the two groups, and thus, *prima facie*, the effects of mask relevance in particular should have been the same in both groups.

Experiment 4

In Experiments 2 and 3, we successively reduced the possibility of perceptual interactions between prime and Stimulus 2; in Experiment 4, we removed the possibility of prior motor-response activation or inhibition affecting the target response systematically. In one group, termed the fixed-mapping group, conditions were the same as in the mixed-family group from the previous experiment. In a second group, termed the variable-mapping group, we manipulated the response mapping on a trial-by-trial basis.

Participants in the fixed-mapping group used keys of a left and a right computer mouse to respond to left targets and right targets, respectively, as in the previous experiments. Participants in the variable-mapping group used the keys with upward and downward pointing arrow on it on a standard computer keyboard. These two keys are arranged on a vertical axis. More important, which of these keys was to be mapped on the response "right" (and which one on the response "left") was determined randomly on a trial-by-trial basis. The current response mapping was indicated by the positions of the words *RECHTS* (*RIGHT*) and *LINKS* (*LEFT*) that appeared above and below the target on the computer screen. Sometimes *RECHTS* appeared above the target and *LINKS* appeared below; in other trials, it was the other way around. The up versus down position of these response labels (relative to the

target) indicated whether the upward or downward pointing arrow key was to be pressed to indicate a “right” response or “left” response. For example, if *RECHTS* appeared below the target and *LINKS* above it, participants were to respond with the downward pointing arrow key if the target required the response “*rechts*.”⁶

The response labels appeared along with target onset and were removed after each response. This means that possible interactions between (a) motor activation and inhibition elicited by events prior to the target and (b) motor activation engendered by the target itself should cancel out via randomization: For a given response mapping, administered in a random half of the trials, whatever response activation or inhibition accrues prior to the target will facilitate the response required for one of the two kinds of target and interfere with the response required for the other kind of target. The pattern of facilitation of and interference with target responses reverses for the other response mapping, administered in the other half of the trials. Costs and benefits of prime-induced response activation and inhibition should thus cancel out. All of the current accounts therefore predict little net effect of the prime on target responses. In contrast, the predictions of the evaluation window account remain unchanged, because NCE is caused at an abstract level of categorizing prime and target into the left versus right categories, before the information on stimulus category is channeled into a specific motor response according to the current response mapping. The prediction was therefore that NCE should remain intact in the group with variable mapping.

Method

Procedures were the same as in Experiment 3 with the following exceptions. Procedures for the fixed-mapping group were the same as for the mixed-family group of Experiment 3. For the variable-mapping group, response mapping was randomly determined from trial to trial. The positions of the response labels *RECHTS* and *LINKS* relative to the target indicated which of two keys on the computer keyboard was to be pressed for a left or right response as described in the introduction to this experiment. The response labels appeared centered horizontally above and below the target, with a vertical distance of 1.3° to the target. The size of the response labels was approximately 0.7° × 1.3°. On any given trial, the labels appeared with target onset, and the labels were taken off the screen as soon as a response was entered. Because we anticipated somewhat longer response latencies in the variable-mapping group than in the fixed-mapping group, the target remained on screen for up to 800 ms in the variable-mapping group, rather than for up to 600 ms, as in the fixed-mapping group. The rules for the performance-contingent payoff were accordingly adapted: Participants in the variable-mapping (fixed-mapping) group received 1 Euro cent (equivalent to U.S. \$1.23) for each correct response within 800 ms (600 ms) after target onset.

Participants were 30 University of Freiburg students with different majors and a few high-school students. Mean age was 23 years, ranging from 17 years to 33 years. One participant’s data were excluded because of empty cells.

Results

Response latencies were preprocessed as in Experiment 1, excluding 4.32% of the trials. Mean correct response latency was 542

ms; mean error rate was 7.65%. In Figure 6, compatibility effects are shown for the fixed-mapping group (left panels) and the variable-mapping group (right panels) as a function of Stimulus 2 category and stimulus family in the latency domain (upper panels) and in the accuracy domain (lower panels).

The compatibility effects were submitted to an analysis of variance with repeated measures on the factors stimulus family, Stimulus 2 category, and between-participants factor group. For the latency domain, none of the effects or interactions was significant (largest $F = 1.83$, smallest $p = .19$). The analogous analysis in the accuracy domain revealed a significant interaction of stimulus family and group, $F(1, 27) = 6.22$, $p = .02$, $\eta_p^2 = .19$. No other effect or interaction was significant (largest $F = 2.36$, smallest $p = .10$). The interaction of stimulus family and group had the following form: In the fixed-mapping group, NCE was larger for trials with letter targets ($M = -2.40\%$) than for trials with arrow targets ($M = -0.05\%$) and vice versa in the variable-mapping group ($M = -0.04\%$ and $M = -3.02\%$, respectively).

Separate t tests confirmed significant overall NCE in the fixed-mapping group and in the variable-mapping group in the latency domain ($M = -9$ ms), $t(13) = -4.34$, $p < .01$, and ($M = -15$ ms), $t(14) = -4.65$, $p < .01$, respectively. NCE in the accuracy domain was $M = -0.84\%$, $t(13) = -2.28$, $p = .04$, and $M = -1.59\%$, $t(14) = -1.94$, $p = .07$, respectively. Analyses of variance of response latencies and percentage-correct scores are presented in the Appendix.

Discussion

Results were clear cut. NCE was found in both groups, irrespective of kind of Stimulus 2. This pattern of results is consistent with the evaluation window account but is difficult to explain by any of the other accounts.

Experiment 5

In Experiments 5 and 6, Stimulus 2 was omitted, and priming was examined in a paradigm with prime and target, without intervening stimulus. According to all of the current accounts of NCE in arrow priming, Stimulus 2 is necessary for NCE to occur, if for different reasons. In fact, when a mask was left out in previous research on arrow priming, PCE was always found as reviewed in the introduction.

As already mentioned, we acknowledge that one component of priming is caused by the interaction and regulation of motor activation elicited by prime and target, and we accept that in the absence of Stimulus 2, PCE contributed to observed compatibility effects via this component. To reduce the impact of this component, we used somewhat less overlearned stimulus–response mappings than in Experiments 1–3. Specifically, stimuli (double arrows and squares with a point in it, see Figure 2) pointed up or down. They were to be responded to with a left key, versus a right key, as before. Removing the spatial correspondence between

⁶ Moving to the vertical dimension for response keys and for the screen positions of response labels avoided pronounced differences between the two intermixed response mappings in terms of spatial compatibility between left and right responses on the one hand and the positions of response keys as well as response labels on the other hand.

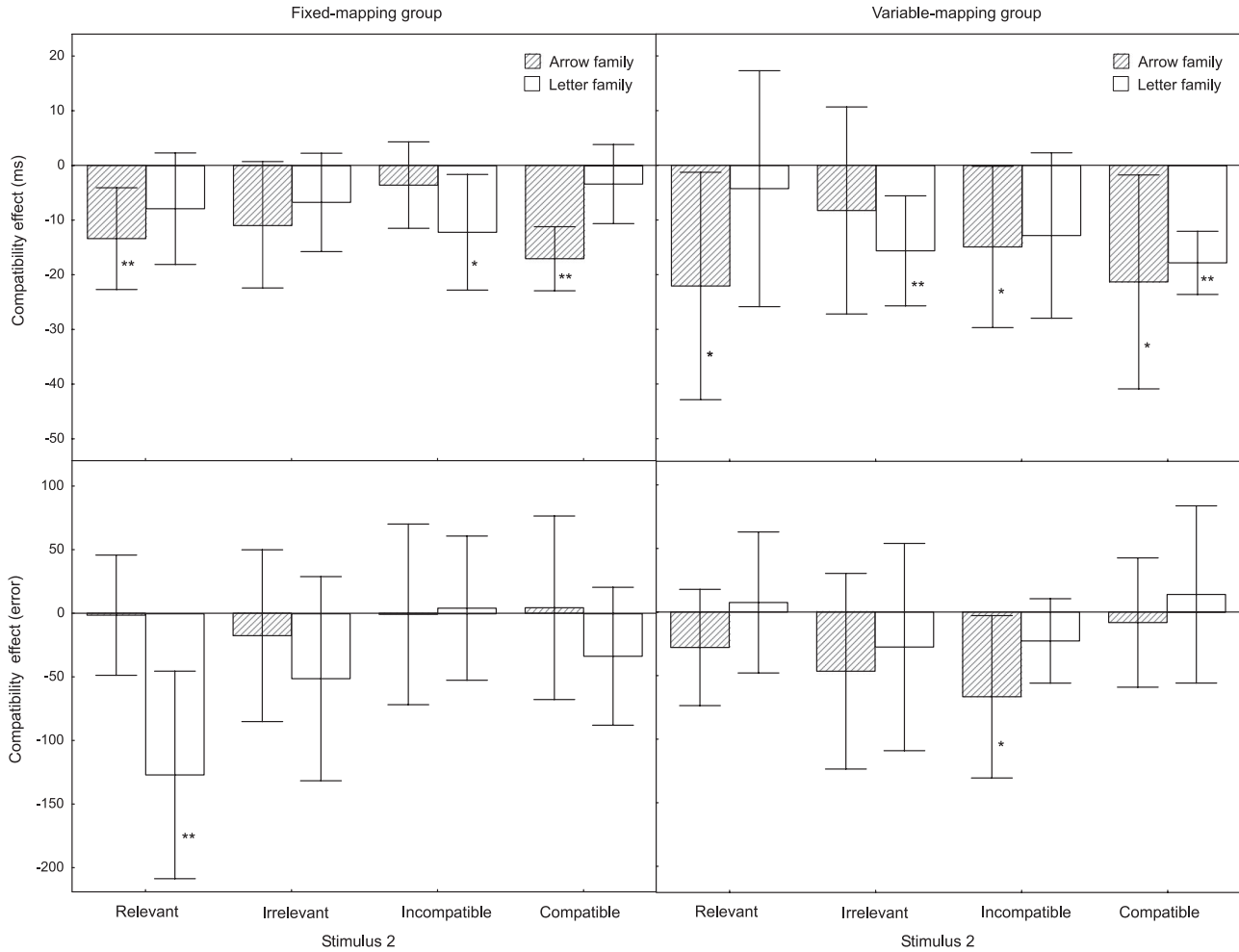


Figure 6. Experiment 4. Compatibility effects in the fixed-mapping group (left panels) and the variable-mapping group (right panels) in the latency domain (upper panels) and in the accuracy domain (lower panels), based on arcsine-transformed values multiplied by 1,000) as a function of Stimulus 2 (relevant, irrelevant, incompatible, compatible) and stimulus family of prime and target (arrow, letter). Error bars show the 95% confidence intervals. * $p < .05$. ** $p < .01$.

stimuli and responses (i.e., moving from the vertical axis for stimuli to a horizontal axis for responses) should somewhat slow down and reduce spontaneous motor activation for the appropriate response (Proctor & Cho, 2006).

In Experiment 5, we implemented a set of conditions that should be favorable for NCE to occur according to the evaluation window account. According to that account, NCE should be most likely for a prime–target SOA of intermediate size. For a short SOA, the onset of the evaluation window is locked to the prime serving as a go signal predicting the immediately impending target onset. As the temporal distance between prime and target is increased, there should come a point at which the prime is no longer perceived as useful for signaling an immediately impending target onset, and the synchronization between prime and window onset should break down. Beyond this point, window onset occurs after prime onset. Thus, as SOA is further increased, PCE should decrease and eventually turn to NCE once the prime is effectively excluded from the evaluation window. As the distance between prime and

target is further increased, NCE should return to zero because prime-driven counter activation is now likely to have decayed prior to window onset.

The adoption of an evaluation window that is no longer locked to the prime should be further encouraged when SOA randomly changes from trial to trial. As stated in Assumption 3, when the temporal parameters of stimulus presentation are repetitive and predictable, it is convenient to use the prime (more precisely, the last stimulus occurring before the target) as a go signal for setting up an evaluation window. Thus, when SOA is fixed or manipulated across blocks, synchronization of the evaluation window with the prime is more strongly supported, fostering inclusive response windows and PCE, than when SOA randomly changes from trial to trial.

In Experiment 5, we implemented these conditions. It was fashioned after Experiment 3 by Klauer et al. (2009) in the evaluative-priming paradigm. In particular, three levels of SOA were administered, with SOA changing randomly on a trial-by-

trial basis. We expected to find PCE for the short SOA, NCE for the medium SOA, and reduced NCE or no priming at the long SOA, and thus, the results should conceptually replicate those of the Klauer et al. (2009) study in the domain of arrow priming.

Method

Participants were 97 University of Freiburg students with different majors; mean age was 23 years, ranging from 18 years to 44 years. One participant was an extreme outlier, according to Tukey, with mean error rate of 63.32% in the total sample's distribution of error rates ($M = 3.71$, $SD = 0.57$). This participant was excluded from the analysis.

The stimuli used in this experiment are shown in Figure 2. There were arrows pointing up or down and squares with a dot at the upper or lower inner border. Stimuli were to be classified as pointing up or down using keys of a left and a right computer mouse. Key assignment was counterbalanced so that the response "up" ("down") was mapped on the key of the right (left) computer mouse for half of the participants, whereas "up" ("down") was mapped on the key of the left (right) mouse for the other half of participants. The response mapping was also indicated on the screen by means of the letters *oo* (signaling the response "*oben*" [up]) or *uu* (signaling the response "*unten*" [down]) placed to the left and right of the screen center with a horizontal distance of 1.5° to the target. The size of the response labels was approximately $0.7^\circ \times 1.2^\circ$. On any given trial, the labels appeared with target onset. They remained on screen until a response was given.

Procedures were otherwise the same as in Experiment 1, except for the following differences. Participants saw only one prime displayed for 40 ms. The SOA between prime and target was randomly chosen for each trial to be one of the following three values: 120 ms, 240 ms, or 360 ms.

One prime or target was always an arrow stimulus, the other one was always a square stimulus. This restriction was imposed to minimize perceptual interactions between primes and targets and, in particular, to avoid repetition of identical stimuli within one trial. Whether the target was an arrow or a square stimulus was counterbalanced across trials. Prime category (up vs. down), target category (up vs. down), and SOA (120 ms, 240 ms, or 360 ms) were crossed orthogonally in each block of trials. There were 16 experimental and two practice block of 48 trials each.

Results

Response latencies were preprocessed as in Experiment 1, excluding 5.76% of the trials. Mean correct response latency was 379 ms; mean error rate was 4.13%.

Compatibility effects are shown in Figure 7 as a function of SOA in the latency domain (upper panel) and in the accuracy domain (lower panel). As can be seen, compatibility effects in the latency data were significantly positive for the short SOA, $t(95) = 4.02$, $p < .01$, significantly negative for the medium SOA, $t(95) = -2.69$, $p < .01$, and more or less absent for the long SOA, $|t| < 1$. There were no individually significant compatibility effects in the accuracy domain, but priming effects moved from a comparatively large PCE at the shortest SOA, $t(95) = 1.69$, $p = .09$, to almost zero compatibility effects for the longer SOAs (both $|t| < 1$).

Analyses of variance with factor SOA revealed that the effect of SOA on compatibility effects was significant in the latency domain, $F(2, 190) = 18.06$, $p < .01$, $\epsilon = .99$, $\eta_p^2 = .16$, and in the accuracy domain, $F(2, 190) = 3.12$, $p = .05$, $\epsilon = .95$, $\eta_p^2 = .03$. Analyses of variance of response latencies and percentage-correct scores as a function of prime category, target category, and SOA are presented in the Appendix.

Discussion

In this experiment, priming effects were examined without mask or other kind of stimulus intervening between prime and target. Procedural parameters were chosen so as to create propitious conditions for the emergence of NCE according to the evaluation window account, the most important choices being (a) to use an intermediate SOA, (b) to make motor activation somewhat less fluent to reduce priming at the motor level, and (c) to vary SOA randomly on a trial-by-trial basis.

The results revealed the expected pattern of effects. NCE was observed for the intermediate SOA, PCE for the short SOA, and there were no compatibility effects at the long SOA. The results thereby mirror the pattern observed by Klauer et al. (2009, Experiment 3) in the domain of evaluative priming.

When no stimulus intervenes between prime and target, PCE is typically observed (e.g., Jaśkowski, 2008; Klapp & Hinkley, 2002; Lleras & Enns, 2004; Verleger et al., 2004). However, in these studies, relatively short and fixed prime-target SOAs were used, and/or motor-response selection was as simple and immediate as in our Experiment 1, setting up conditions that favor PCE according to the above discussion.

NCE without mask is in line with the evaluation window account but is difficult to explain in terms of self-inhibition, object updating, or mask-triggered inhibition, simply because no stimulus whatsoever intervened between prime and target. Again, it may be possible to add auxiliary assumptions enabling the accounts by object updating and mask-triggered inhibition to account for the present pattern of compatibility effects post hoc. In particular, under both accounts, it could be argued that target onset itself triggers motor suppression. This suppression, however, may not be strong enough to overcome and reverse the motor activation elicited by the prime under the short SOA, so that PCE is observed. At the intermediate SOA, it might be assumed that motor activation induced by the prime has already decayed to a level at which motor suppression triggered by target onset can successfully overcome and suppress the prime-induced motor activation, leading to NCE. Finally, at the long SOA, the prime-induced motor activation has effectively decayed to a baseline level of no activation. If so, there is no motor activation to be suppressed any more.

This post hoc explanation would be rendered even less plausible, however, if NCE could be demonstrated even for shorter SOA such as the 120 ms SOA for which PCE was observed in the present experiment.

According to the evaluation window account, SOA per se is not important (within limits) for determining whether NCE or PCE is obtained. Rather, the effects of SOA are assumed to be mediated by the temporal positioning of the evaluation window. As a consequence, it should be possible to obtain NCE even with shorter a SOA, such as the 120 ms SOA for which PCE was observed in the present experiment. This was the objective of Experiment 6.

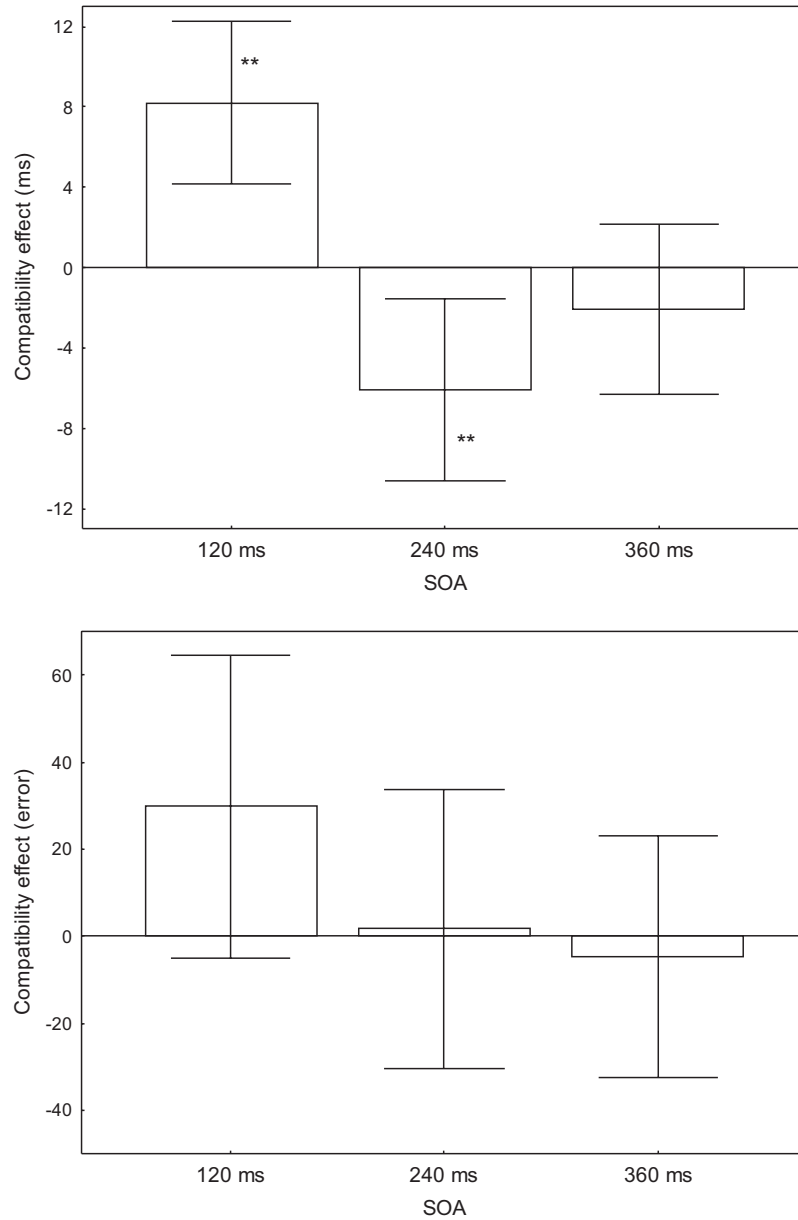


Figure 7. Experiment 5. Compatibility effects in the latency domain (upper panel) and in the accuracy domain (lower panel, based on arcsine-transformed values multiplied by 1,000) as a function of SOA (120 ms, 240 ms, 360 ms). Error bars show the 95% confidence intervals. SOA = stimulus-onset asynchrony. * $p < .05$. ** $p < .01$.

Experiment 6

In Experiment 6, a prime–target SOA of 120 ms was used, as in Experiment 1 and in the second group of Experiment 2, an SOA that is also representative for many previous studies in the masked-arrow paradigm. According to the evaluation window account, one way to achieve NCE with an SOA this short is to present a stimulus between prime and target that unambiguously partitions the stream of stimuli into separate perceptual episodes. Another possibility is to create conditions under which the prime is not used as a go signal signaling the impending onset of the target and under

which the evaluation window is therefore unlikely to start much before target onset.

In Experiment 6, we attempted to achieve this by removing the predictive value of the prime for target onset so that participants were taken by surprise by target onset. Participants saw a stream of several stimuli in any given trial. Across trials, stream length was randomly varied between two and nine stimuli. Participants were asked to respond to the last stimulus of each stream. We were interested in the effect of the penultimate stimulus, referred to as the prime in the following, on the response to the last stimulus, referred to as the target in the following.

Because stream length is randomly varied, the prime is no longer predictive of target onset (except for, to be precise, in the longest stream). It is difficult to say how participants position the evaluation window in this situation. One strategy would be to bet on a stream of average length and to open the window as soon as, or somewhat before, an average number of stimuli has been presented, implying that the evaluation window has to be reset if stream length is longer than expected. Another strategy is to wait until the target is identified (as the last stimulus to appear) and to open the evaluation window only then. Generally speaking, the temporal positioning of the evaluation window is likely to show more variability within participants in this situation. Klauer et al. (2009) argued, however, that target onset would catch most participants by surprise for the shortest streams, so that participants would not have opened an evaluation window before target onset for these streams. If so, the evaluation window should exclude most of the prime-derived activation, and NCE should occur. We thus predict NCE for the shortest streams. In fact, in the context of evaluative priming, this was just what was observed (Klauer et al., 2009, Experiment 4).

In Experiment 6, we again effectively eliminated the contribution of the motor level to compatibility effects through the use of a variable response mapping as in Experiment 4. This was done because we expect prime-derived motor activation to contribute a strong and overriding PCE component when SOA is short and its impact is left unchecked.

Method

Participants were 105 University of Freiburg students with different majors; mean age was 24 years, ranging from 18 years to 37 years. Three participant were extreme outliers, according to Tukey, with mean error rates of 49.45%, 50.37%, and 53.05% in the total sample's distribution of mean error rates rate ($M = 6.69\%$, $SD = 0.64$); one more participant was an extreme outlier, according to Tukey, with mean reaction time data of 988 ms in the total sample's distribution of mean response latencies ($M = 659$ ms, $SD = 86$). These participants were excluded from the analysis.

The stimuli used in this experiment were the ones already used in Experiment 5 (see Figure 2). In a trial of Experiment 6, participants saw a stream of stimuli. Stream length was randomly sampled for each trial, with a range from two to nine stimuli and with the restriction that each length occurred equally often within each block of trials. We refer to the last stimulus of each stream as the target and to the penultimate stimulus as the prime. Stimulus duration for all stimuli other than the target was 60 ms. The SOA of successive stimuli was 120 ms. Note that this SOA is identical to the short SOA in Experiment 5, for which PCE resulted.

Successive stimuli always changed from being arrows to squares or vice versa (see Figure 2) to reduce perceptual interactions among the stimuli of a stream, in particular to avoid the presentation of two or more identical stimuli in succession. The up versus down category of these stimuli was determined randomly. Response assignments were randomly set for each trial. The participant saw, as in Experiment 5, the letters *oo* (for *oben* [up]) and *uu* (for *unten* [down]) to the left or the right of the target. The left versus right position of these letters indicated whether the left or

the right mouse key should be pressed to give an up or down response, respectively. For example, if the letter *o* appeared left of the target and *uu* appeared right of the target, the left key was to be pressed for a target requiring the response "*oben*" and the right key was to be pressed for a target requiring the response "*unten*." The left–right position of the letters *oo* or *uu* was randomly chosen for each trial. These response labels appeared along with the target, and they stayed on screen until a response was registered. Participants could identify targets as such because of the flanking response labels and because the target was the last stimulus to appear.

The response window within which a correct response contributed to the participants' payoff was 800 ms, and the target was taken off the screen after 800 ms (or as soon as a response was registered whichever event occurred earlier). There were 16 experimental blocks of 48 trials each, preceded by two practice blocks of 48 trials each.

Results

Correct response latencies were preprocessed as in Experiment 1, excluding 3.09% of the trials. Mean response latency was 664 ms; mean error rate was 5.81%. As in Klauer et al.'s (2009) Experiment 4, stream lengths were analyzed in four bins (Bin 1: stream lengths two and three; Bin 2: stream lengths four and five; Bin 3: stream lengths six and seven; Bin 4: stream lengths eight and nine).

Compatibility effects are shown in Figure 8 as a function of stream length in the latency domain (upper panel) and the accuracy domain (lower panel). As can be seen, significant NCE was found for the shortest streams in the latency data, $t(100) = -2.43$, $p = .02$. PCE prevailed for longer streams, smallest $t(100) = 2.76$, largest $p = .01$. In the accuracy data, only PCE for stream lengths six and seven was individually significant, $t(100) = 2.11$, $p = .04$.

Analyses of variance of the compatibility effects with factor stream length (four bins; within-participants) showed a main effect of stream length, $F(3, 300) = 8.70$, $p < .01$, $\epsilon = .95$, $\eta_p^2 = .08$, in the latency domain but no significant effect in the accuracy domain, $F(3, 300) = 1.80$, $p = .15$, $\epsilon = .98$, $\eta_p^2 = .02$. Analyses of variance of response latencies and percentage-correct scores as a function of prime category, target category, and stream length are presented in the Appendix.

Discussion

In Experiment 6, another unique prediction of the evaluation window account was tested. By presenting streams of stimuli of randomly varying lengths, the predictive value of the prime as a cue signaling the impending onset of the target was largely removed. This should make it impossible to use the prime as a go signal for setting up an early evaluation window. It is an open question how the evaluation window is positioned temporally in this situation, but it is plausible that target onset would catch most participants by surprise for the shortest streams, given that most streams were longer. Consequently, it is unlikely that window onset could be positioned much before target onset for such streams. This leads to exclusive evaluation windows, excluding the prime, for short streams and to the prediction that NCE should occur for such streams. Replicating results obtained in the

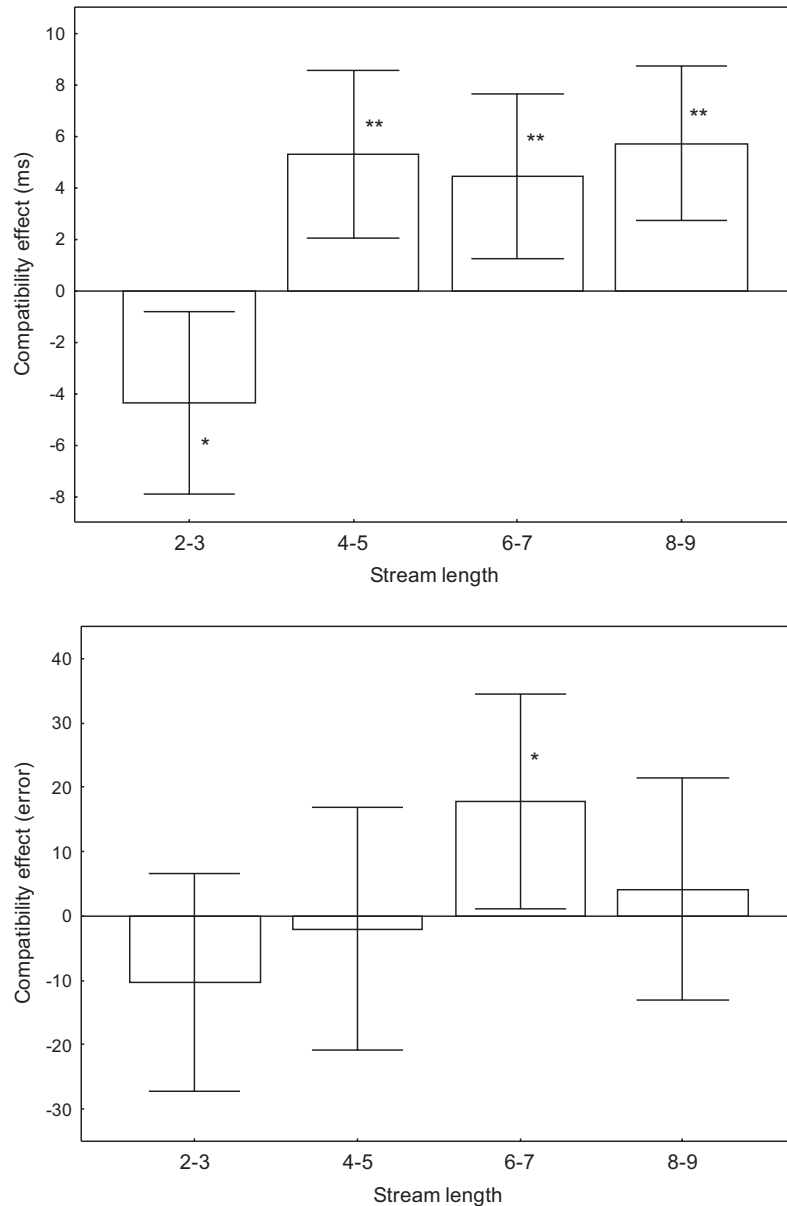


Figure 8. Experiment 6. Compatibility effects in the latency domain (upper panel) and in the accuracy domain (lower panel, based on arcsine-transformed values multiplied by 1,000) as a function of stream lengths (2–3, 4–5, 6–7, 8–9). Error bars show the 95% confidence intervals. * $p < .05$. ** $p < .01$.

evaluative-priming paradigm (Klauer et al., 2009), NCE in fact emerged for the shortest streams.

As in Experiment 4 of the present series, response mapping was randomly switched on a trial-by-trial basis to eliminate a motor-level component of compatibility effects. The use of variable response mappings typically leaves priming effects intact (Abrams, Klinger, & Greenwald, 2002; Klauer, Mierke, & Musch, 2003, Experiment 3A), and it did not eliminate compatibility effects in the present experiment. As in Experiment 4, this again suggests that a more central stage of categorizing stimuli contributes to compatibility effects as postulated by the evaluation window account.

General Discussion

A series of six experiments was presented. The experiments fell into two groups. In Experiments 1–4, primes and targets were separated by an intervening stimulus, Stimulus 2.

In Experiment 1, we replicated NCE in arrow priming using SOAs of 60 ms between successive stimuli. A new feature was that both prime and Stimulus 2 were sometimes response-valent stimuli themselves, that is, double arrows pointing left or right. Stimulus 2 was a relevant mask, an irrelevant mask, or a response-valent stimulus. NCE was largest in trials with relevant masks and smaller in trials with irrelevant masks, whereas there was mostly

PCE for trials with response-valent Stimulus 2. The evaluation window account cannot explain the effects of mask relevance that we attribute to perceptual interactions of prime and mask as per object updating or that we attribute to motor-control processes as per mask-triggered inhibition.

In Experiment 2, these results were replicated in a first group of participants. In an additional group of participants, stimulus duration and SOA were lengthened by a factor of 3 that should reduce perceptual interactions between prime and Stimulus 2. As predicted by the evaluation window account, this leveled the effect of mask relevance on NCE and resulted in NCE for trials with response-valent Stimuli 2. NCE for response-valent Stimuli 2, however, was still significantly smaller than NCE in trials with masks, so that Stimulus 2 effects on NCE were reduced but not eliminated.

In Experiment 3, we further reduced the potential for perceptual interactions of primes and Stimulus 2 by using perceptually dissimilar primes and Stimuli 2. Whereas effects of Stimulus 2 on NCE were replicated in a group with perceptually similar primes and Stimuli 2 as in Experiment 2, all such effects were eliminated in a second group in which primes and Stimuli 2 were dissimilar, as predicted by the evaluation window account and consistent with the account by object updating.

In Experiment 4, response mapping was varied randomly. Specifically, two different mappings are possible in mapping a left versus right response on two response keys. For participants of a fixed-mapping group, response mapping was held constant throughout the experimental session. For participants of a variable-mapping group, response mapping was randomly selected for each trial. Nevertheless, NCE occurred in both groups, and there were no effects of Stimulus 2 on NCE. This is consistent with the evaluation window account but difficult to explain by the current accounts that attribute NCE to the effects of motor activation and/or suppression elicited by events prior to target onset (Boy & Sumner, *in press*).

In Experiments 5 and 6, Stimulus 2 was left out altogether. None of the current accounts can account for NCE without Stimulus 2. The evaluation window account does predict NCE under certain conditions. Two such sets of conditions were implemented in Experiment 5 (intermediate SOA) and Experiment 6 (target onset comes as a surprise). In Experiment 5, we manipulated prime–target SOA in three steps, 120 ms, 240 ms, and 360 ms, with SOA selected randomly from these three levels for each trial. This resulted in PCE for the short SOA, NCE for the medium SOA, and little priming of any kind for the long SOA. In Experiment 6, we presented streams of stimuli of varying length and asked participants to respond to the last stimulus of each stream, the target. SOA between two successive stimuli was 120 ms. The stimulus immediately preceding the target caused NCE in the shortest streams (*i.e.*, for streams of lengths 2 and 3) as predicted.

The Evaluation Window Account Explains Some, But not All Effects in Arrow Priming

Research on arrow priming has generated a rich set of findings, suggesting that two broad classes of processes shape compatibility effects observed in that paradigm: (a) regulation and inhibitory control of motor activation elicited by primes and targets and (b)

perceptual interactions of prime and Stimulus 2. One purpose of the present research was to propose that a third class of processes might be involved, operating at a more central stage of categorizing stimuli into the task-relevant categories.

The evaluation window account, considered in isolation, is consistent with a number of effects found in arrow priming, but it cannot account for others. First of all, it accounts for the basic NCE result with the assumption that Stimulus 2 comes to serve as a go signal signaling the impending onset of the target and that the onset of the evaluation window will therefore be locked to Stimulus 2 onset, excluding the prime. It is also broadly consistent with the temporal parameters associated with NCE: PCE instead of NCE occurs for short prime–target SOAs (Kiesel *et al.*, 2008; Lleras & Enns, 2005; Schlaghecken & Eimer, 2000, 2002). This boundary condition would be attributed to the difficulty of positioning the evaluation window with such precision as to exclude the prime reliably when prime–target SOA is short, and instead, prime-derived activation will be included in the interval (remember that we assume that window onset will usually occur anticipatorily before target onset to ensure that full use is made of target-derived activation).

Related to the above, PCE reliably occurs when there is no Stimulus 2 (*e.g.*, Jaśkowski, 2008; Klapp & Hinkley, 2002; Lleras & Enns, 2004; Verleger *et al.*, 2004). According to the evaluation window account, PCE flows from the implication that the onset of the evaluation window would tend to be locked to the prime when there is no Stimulus 2. In a similar vein, Boy *et al.* (2008) compared conditions in which Stimulus 2 intervened between prime and target with conditions in which Stimulus 2 was present before the prime, but at a different location, and was then gradually moved to mask the prime at that point in time at which it appeared in the first condition. NCE was observed in the first condition, and PCE was observed in the second condition, in line with the fact that Stimulus 2 is less useful as a signal of impending target onset when it appears before the prime. Again, the prime itself is then more likely to be used as such a signal to which the onset of the evaluation window is locked, leading to PCE.

Other effects are less readily explained by the evaluation window account. Schlaghecken and Eimer (2000; Schlaghecken & Eimer, 2002) discovered a so-called central-peripheral asymmetry. When prime and Stimulus 2 are presented peripherally and the target is presented centrally, PCE occurred instead of NCE. The evaluation window account needs additional assumptions to account for this finding. For example, it could be assumed that the distinction between prime and Stimulus 2 is perceptually blurred when these stimuli are presented out of focus, making it more difficult to use Stimulus 2 to partition the stream of incoming stimuli into separate perceptual episodes. As a consequence, the evaluation window might become locked to the onset of the prime–Stimulus 2 compound rather than to Stimulus 2 onset.

Similarly, a number of experiments have shown more NCEs when primes are less discriminable and more PCEs when primes are more discriminable (*e.g.*, Eimer & Schlaghecken, 2002; Klapp, 2005; Klapp & Hinkley, 2002; Lleras & Enns, 2004, 2005; Schlaghecken & Eimer, 2006; Sumner, Tsai, Yu, & Nachev, 2006). There is some debate as to whether this negative relationship between prime strength and NCE reflects a causal role of sensory prime strength or confounded factors such as

differences between the masks used to manipulate prime strength (Jaśkowski & Verleger, 2007; Schlaghecken, Blagrove, & Maylor, 2008; Sumner, 2007). But if it is indeed related to the strength of the sensory prime representation, it is difficult to see how the evaluation window account, as a stand-alone account, could explain it. Weaker primes should simply cause smaller compatibility effects, smaller NCE when excluded from the evaluation window, and smaller PCE, when included in that window, than stronger primes. On the other hand, an antagonistic tendency might be that Stimulus 2 is relatively more prominent and more likely to be used as a go signal than the prime to the extent to which the prime is weak.

As already mentioned, effects of the nature of Stimulus 2 should be limited according to the evaluation window account. In particular, for Stimuli 2 sufficiently distinct from primes and targets to be effective in partitioning the stream of incoming sensory information into distinct perceptual episodes, it should make little difference whether Stimulus 2 is a relevant or an irrelevant mask and whether it is a mask or one of the stimuli used as primes and targets. But these factors, mask relevance and whether Stimulus 2 itself is response valent, have a major impact; effects of mask relevance are well documented in the literature as already reviewed and were replicated in the present studies (Experiments 1 and 2); effects of whether Stimulus 2 itself is response valent were also found in our studies (Experiments 1, 2, and 3).

For these reasons, the evaluation window account is seen as describing only one of several mechanisms shaping PCE and NCE in arrow priming. It is a mechanism that is new and that has hitherto not been considered. Consequently, we focused on showing in Experiments 1–4 that even when factors currently considered obligatory for the occurrence of NCE in arrow priming are successively removed, NCE as predicted by the evaluation window account can still be observed.

Across these experiments, NCE tended to decrease in size. For example, considering trials with masks, mean NCE for Experiments 1–4 was, respectively, -22 ms, -28 ms, -16 ms, and -11 ms. This is in line with the assumption that we removed components of the NCE effect that reflect perceptual interactions of prime and mask as well as inhibitory control of motor activation. In this view, the residual NCE reflects the operation of the mechanism specified in the evaluation window account. Note also that NCE was surprisingly robust, across these experiments, over a wide range of mean latencies and error rates.

In Experiments 5 and 6, we tested further new predictions of the evaluation window account. According to that account, NCE can be obtained even when no stimulus intervenes between prime and mask. Again, none of the current accounts predicts NCE without intervening stimulus. Nevertheless, small but significant NCEs were observed in both experiments in the conditions under which NCE is predicted by the evaluation window account. What is more, these experiments conceptually replicate analogous findings for evaluative priming, increasing one's confidence in the replicability of these effects.

The Motor Locus of Prime–Target Compatibility Effects

One issue raised by these experiments concerns the locus of NCE and PCE in arrow priming. The evaluation window account

claims that compatibility effects are in part mediated by categorizations of the incoming stimuli into the task-relevant categories. We concur that a major contribution to NCE in arrow priming is located at the motor level. What is the evidence for a motor locus of compatibility effects in arrow priming, and does it rule out a contribution caused at the categorization level?

NCE in arrow priming is reflected in lateralized readiness potentials (LRP), an electrophysiological measure of response activation (Eimer & Schlaghecken, 1998, 2003; Praamstra & Seiss, 2005). These data show an initial deflection associated with the prime, followed by a reversal consistent with inhibition and last a strong deflection associated with the response. In behavioral data, a further oscillation has recently been reported at longer SOAs so that following the NCE, there was a small rebound back to PCE (Sumner & Brandwood, 2008). But as pointed out by Eimer, Schubö, and Schlaghecken (2002), one cannot necessarily conclude that “those effects operate solely at motor levels. Inhibition in the masked prime paradigm may be located at perceptual or central stages and may only subsequently affect motor activation processes” (p. 4).

Similarly, the negative bias induced by primes also occurs with free-choice responses; that is, there is a preference to choose the unprimed response when responses can be freely chosen (Klapp & Haas, 2005; Klapp & Hinkley, 2002; Schlaghecken, Klapp, & Maylor, 2009), again suggesting that motor inhibition of the primed response is involved. But free choice might be triggered by detecting random increases in counters operating at the categorization stage with the detection of increases impaired for the primed counter as per the evaluation window account. Eimer et al. (2002) and Schlaghecken et al. (2009) each reported evidence suggesting a motor-level locus of NCE in arrow priming along with conflicting evidence suggesting the involvement of processes at more abstract levels not tied to specific motor responses.

Some of the strongest evidence for a motor-level locus has been presented by Boy and Sumner (in press). In their experiments, participants learned arbitrary associations between stimuli (vertical and horizontal lines) and responses (left versus right button presses). In one condition of each experiment, SOA between masked primes and targets was 40 ms; in a second condition, it was 150 ms. PCE emerged in the first condition; NCE emerged in the latter, as the stimulus–response association was learned. It is important to note that when the response mapping was switched in the middle of the experiment (Experiments 3 and 4), NCE reversed to PCE of approximately equal size and vice versa (see in particular the fine-grained analysis in Boy and Sumner's Experiment 4). Because switching the response mapping should not affect the categorization level (i.e., stimuli still have to be classified according to whether they are horizontal or vertical lines) or earlier perceptual analysis, this is strong evidence for an involvement of the motor stage. Another possibility, however, is that participants use abstract left versus right categories and learn to interpret features of horizontal and vertical lines as evidence for the left and right categories. If so, the evaluation window account would predict the same effect of switching the response mapping because the stimulus to category mapping has to be relearned if the response mapping is switched. In this interpretation, the effect of switching the response mapping would be located at a more abstract locus than that of motoric button presses.

A similar idea was used in the present Experiments 4 and 6: Response mapping was switched not only once but randomly on a trial-by-trial basis with the currently active response mapping signaled only along with the target. As explained in the introduction to Experiment 4, a contribution at the level of motor responses in terms of key presses to NCE should thereby be eliminated via randomization. A substantial amount of NCE was nevertheless obtained, suggesting that its locus must be at a level that is more abstract.

Boy and Sumner (in press) further noted that if NCE reflects a perceptual effect, it should be present in all blocks in their experiments (in the 150 ms SOA conditions), but as already mentioned, PCE and NCE built up gradually and in parallel as the arbitrary stimulus to response category mapping was learned. Some amount of learning is also implied by the evaluation window account, whatever the categories that are used: Participants have to tune into the repetitive and predictable nature of stimulus presentation before they can effectively use the mask to position the evaluation window. For this reason, a gradual emergence of PCE and NCE is compatible with the evaluation window account.

Taken together, the evidence for an involvement of the motor level in the generation of observed NCEs is relatively strong, but the evidence for the absence of a contribution at more central levels is ambiguous. At the same time, the present experiments provide evidence for the hypothesis that a more abstract level than the level of motor control for the required key press responses is involved.

Models of Evidence Accumulation and Huber's (2008) Habituation Model

In this section, we consider the relationship of the evaluation window account to general-purpose models of evidence accumulation in binary decisions and to Huber's (2008) habituation model covering a variety of inhibitory phenomena in priming and masking.

The evaluation window account incorporates the idea of evidence accumulation in binary decisions that is central to classical models of stochastic information accumulation (Townsend & Ashby, 1983) such as Ratcliff's (1978) diffusion model. Within this class of models, it has, perhaps, the greatest overlap with Usher and McClelland's (2001) so-called leaky, competing accumulator model. Like in that model, we postulate a separate counter for each decision alternative, with lateral inhibition between the two counters (see Footnote 4) and decay of accumulated information over time.

One difference between the present conceptualization and these other models is that the decision is based not on the absolute levels of counter states (e.g., on which counter exceeds a preset decision threshold first) but on increases in counter states. But an accumulation of evidence in counters can be set up to cause an accumulation of evidence in higher level counters sensitive to increases in the lower level counters. It seems likely that decisions based on the absolute levels of these higher level counters would lead to a model that is both consistent with the present conceptualization and very similar in empirical predictions to the classical models just mentioned. A second major difference is that the sensitivity of these higher level counters for detecting increases is assumed to depend on the initial state of the respective lower level counters at

the onset of the evaluation window. It is beyond the scope of the present model to explore whether a formal model incorporating this modification can cope with the same range of data that classical models of evidence accumulation aim to explain (such as the shape of response-latency distributions for correct and false response-latency distributions).

Huber (2008) developed a neural habituation model in the context of word identification experiments. The model was applied to a wide range of priming and interference phenomena. Like the evaluation window account, in this model, it is assumed that NCEs are contributed to at different levels of processing and that the partitioning of the rapidly changing stream of incoming events plays a crucial role in reducing source confusion in evidence accumulation. The major idea is that overexposure to a stimulus results in depletion following initial activation at whatever level of analysis is considered (e.g., visual, orthographic, or semantic in the case of word stimuli). Depletion is comparatively more enduring than activation and produces inhibitory aftereffects such as NCEs. This gives rise to PCE for stimuli presented briefly (only activation is triggered), but NCE is expected for primes presented sufficiently long for depletion to occur.

Applied to arrow priming, the account is not consistent with the finding that NCE is observed for masked primes, whereas PCE occurs for unmasked primes. In fact, as acknowledged by Huber (2008), the habituation model predicts the opposite effect pattern because unmasked primes are more likely to trigger depletion than masked primes. Similarly, a transition from NCE to PCE as prime discriminability increases presents a puzzle to the account by habituation.

Huber (2008) proposed an account of the effects of mask relevance in terms of the habituation model as follows: Consider a brief presentation of a right-pointing prime arrow (presented for P ms), followed by a mask (for M ms) that consists of superimposed left-pointing and right-pointing arrows. In this situation, detection of the right-pointing arrow receives $P + M$ ms of perceptual support and that of the left-pointing arrow receives M ms of support. If $P + M$ ms is sufficient to induce depletion, then there will be more depletion for a right-pointing arrow than for a left-pointing arrow, resulting in NCE in trials with relevant mask.

Because depletion is slower to build up and lasts longer than initial activation, this idea leads one to expect stronger effects rather than diminished effects of mask relevance, if anything, when stimulus durations are increased as in the present Experiment 2. It is therefore difficult to see how the leveling of the effect of mask relevance in the slow group of that experiment could be dealt with by the account of mask relevance derived from the habituation model.

Given these difficulties for the habituation model in accounting for major experimental findings from arrow priming (see also Boy & Sumner, in press) and for the present results, more direct evidence would probably be needed in order to demonstrate that the mechanisms specified by that model contribute to shaping PCE and NCE in arrow priming.

Scope of the Evaluation Window Account

Having first developed the evaluation window account for evaluative priming, here we tested arrow priming as a first step in investigating the scope of the model. Arrow priming seemed a

good choice for several reasons: NCE in arrow priming is to our knowledge the best studied case of NCE, with a rich empirical and theoretical literature to build on. In addition, as a priming paradigm it is perhaps as dissimilar to evaluative priming as possible: In arrow priming, there are only a few stimuli, most often only one per response category, whereas there are typically many stimuli per category in evaluative priming. This has many implications, two noticeable ones being (a) that response latencies are typically substantially larger in evaluative priming than in arrow priming and (b) that prime and target are often the same stimuli in arrow priming, whereas stimulus repetition is rare and most often explicitly ruled out in evaluative priming. Furthermore, words are typically used that have to be analyzed semantically in evaluative priming, whereas arrow priming typically relies on pictorial stimuli with spatial features that are mapped on responses congruent with the stimuli in the spatial features (e.g., the left-pointing arrow is mapped on a left key). Finally, the prime is typically masked in arrow priming, whereas most often there is no mask (other than the target itself) in evaluative priming (Klauer & Musch, 2003).

Despite these many dissimilarities, the present results suggest that the evaluation window account also contributes to shaping PCE and NCE in arrow priming. This is a first step in raising confidence in the contention that the evaluation window account applies to the very wide range of different category-priming paradigms that are studied in different fields of psychology as outlined in the introduction.

The evaluation window account thus may be relevant for many different lines of research in psychology sharing the common use of a category-priming paradigm. One counterintuitive implication of the evaluation window account is that PCE is in part created rather than revealed by priming paradigms. According to the evaluation window account, both PCE and NCE rely in large part on the repetitive nature of priming paradigms in which many trials are presented with the same fixed temporal pattern of events. We assume that participants capitalize on such repetitive structures to position the evaluation window. In particular, window onset will frequently be locked to the prime that in most priming studies reliably predicts the impending onset of the target. When the predictive value of the prime is removed, however, window onset is likely to show more variability, reducing PCE. Moreover, when target onset is not only unpredictable but even comes as a surprise, there is little likelihood that the evaluation window is opened much prior to target onset, and NCE results (see Experiment 6 and Klauer et al., 2009). In natural environments, unpredictability is likely to be the norm rather than the exception, suggesting that the generalizability of findings from category-priming paradigms to situations outside the laboratory may be enhanced by factoring in the predictions of the evaluation window account.

In concluding, let us point to a couple of fields of inquiry outside the area of category priming to which key principles of the evaluation window account have been applied or may be applicable. As in Huber's (2008) neural habituation model, the mechanism described by the evaluation window account may operate at different levels. Consider, for example, the phenomenon of repetition blindness. It refers to the fact that participants are less likely to detect the repetition of a target word in a rapid visual stream than they are to detect a second, different target (Kanwisher & Potter, 1990). Hochhaus and Johnston (1996) proposed that there is a counter for each possible candidate word and that word identifica-

tion relies on determining which counter's activation level has increased the most from the target exposure. In line with the Weber-Fechner law, Hochhaus and Johnston argued that the difference in activation levels between zero and one prior presentation as in the control condition without repetition is easier to detect than the difference between one prior presentation and two prior presentations as for repeated targets. This is consistent with a version of the evaluation window account operating at the level of stimulus identification with a separate counter for each candidate stimulus that can occur.

Similarly, Eder (2006) applied key principles of the evaluation window account to explain the phenomenon of blindness to response-compatible stimuli (Müsseler & Hommel, 1997). The finding is that categorization of a stimulus is impaired if it shares a task-relevant feature with a response that is executed in parallel. In Müsseler and Hommel's (1997) paradigm, participants are instructed to prepare either a left movement or a right movement (e.g., of a joystick). As soon as they are ready to execute the movement, they indicate their readiness to do so and then execute the movement as fast as possible. Simultaneously with the execution, a left-pointing arrow or right-pointing arrow is presented, and participants are to classify it as pointing left or right. The classification is impaired if the left or right movement shares the left or right feature of the arrow. For example, classification of a left-pointing arrow as pointing to the left is impaired relative to the classification of a right-pointing arrow as pointing to the right during the execution of a movement to the left. In the current terminology, the self-generated readiness signal may act as a go signal for opening an evaluation window that thereby excludes activation of the left or right feature stemming from the prior movement-planning phase.

These examples suggest that key principles of the present account may be applicable to account for a range of inhibitory phenomena that unfold in time (e.g., repetition blindness, blindness to response-compatible stimuli) beyond category priming. The key principles are that participants evaluate incoming evidence across a time window, and decisions about stimulus category or identity are driven by changes in evidence weighted according to the Weber-Fechner law. Inhibition by compatible context stimuli results whenever the stream of events supports a clear temporal separation of targets from context stimuli so that context stimuli are excluded from the evaluation window.

References

- Abrams, R. L., Klinger, M. R., & Greenwald, A. G. (2002). Subliminal words activate semantic categories (not automated motor responses). *Psychonomic Bulletin & Review*, *9*, 100–106.
- Bennett, J. D., Lleras, A., Oriet, C., & Enns, J. T. (2007). A negative compatibility effect in priming of emotional faces. *Psychonomic Bulletin & Review*, *14*, 908–912.
- Botvinick, M., Braver, T., Barch, D., Carter, C., & Cohen, J. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.
- Boy, F., Clarke, K., & Sumner, P. (2008). Mask stimulus triggers inhibition in subliminal visuomotor priming. *Experimental Brain Research*, *190*, 111–116.
- Boy, F., & Sumner, P. (in press). Tight coupling between positive and reversed priming in the masked prime paradigm. *Journal of Experimental Psychology: Human Perception and Performance*.
- Clark-Carter, D. (2004). *Quantitative psychological research: A student's handbook*. New York, NY: Psychology Press.

- Degner, J., & Wentura, D. (2010). Automatic prejudice in childhood and early adolescence. *Journal of Personality and Social Psychology*, 98, 356–374.
- Dehaene, S., Naccache, L., LeClec'H, G., Koechlin, E., Mueller, M., van de Moortele, P., & Le Bihan, D. (1998, October). Imaging unconscious semantic priming. *Nature*, 395, 597–600.
- De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Erlbaum.
- Eder, A. (2006). "Common Coding" von Handlung und Bewertung [Common coding of action and valence]. Berlin, Germany: Logos.
- 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, 1737–1747.
- Eimer, M., & Schlaghecken, F. (2002). Links between conscious awareness and response inhibition: Evidence from masked priming. *Psychonomic Bulletin & Review*, 9, 514–520.
- Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology*, 64, 7–26.
- Eimer, M., Schubö, A., & Schlaghecken, F. (2002). The locus of inhibition in the masked priming of response alternatives. *Journal of Motor Behavior*, 34, 3–10.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229–238.
- Fockenberg, D. A., Koole, S. L., & Semin, G. R. (2008). Priming in concert: Assimilation and contrast with multiple affective and gender primes. *Social Cognition*, 26, 647–669.
- Forster, K. I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 3–38). New York, NY: Psychology Press.
- Gawronski, B., Deutsch, R., & Seidel, O. (2005). Contextual influences on implicit evaluation: A test of additive versus contrastive effects of evaluative context stimuli in affective priming. *Personality and Social Psychology Bulletin*, 31, 1226–1236.
- Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996, September). Three cognitive markers of unconscious semantic activation. *Science*, 273, 1699–1702.
- Hochhaus, L., & Johnston, J. C. (1996). Perceptual repetition blindness effects. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 355–366.
- Huber, D. E. (2008). Immediate priming and cognitive aftereffects. *Journal of Experimental Psychology: General*, 137, 324–347.
- Jaśkowski, P. (2007). The effect of nonmasking distractors on the priming of motor responses. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 456–468.
- Jaśkowski, P. (2008). The negative compatibility effect with nonmasking flankers: A case for mask-triggered inhibition hypothesis. *Consciousness and Cognition*, 17, 765–777.
- Jaśkowski, P., & Przekoracka-Krawczyk, A. (2005). On the role of mask structure in subliminal priming. *Acta Neurobiologiae Experimentalis*, 65, 409–417.
- Jaśkowski, P., & Verleger, R. (2007). What determines the direction of subliminal priming? *Advances in Cognitive Psychology*, 3, 181–192.
- Kanwisher, N. G., & Potter, M. C. (1990). Repetition blindness: Levels of processing. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 30–47.
- Kiesel, A., Berner, M. P., & Kunde, W. (2008). Negative congruency effects: A test of the inhibition account. *Consciousness and Cognition*, 17, 1–21.
- Kiesel, A., Kunde, W., Pohl, C., & Hoffmann, J. (2006). Priming from novel masked stimuli depends on target set size. *Advances in Cognitive Psychology*, 2, 37–45.
- Klapp, S. T. (2005). Two versions of the negative compatibility effect: Comment on Lleras and Enns (2004). *Journal of Experimental Psychology: General*, 134, 431–435.
- Klapp, S. T., & Haas, B. W. (2005). Nonconscious influence of masked stimuli on response selection is limited to concrete stimulus-response associations. *Journal for Experimental Psychology: Human Perception and Performance*, 31, 193–209.
- 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, 255–269.
- Klauer, K. C., Eder, A., Greenwald, A. G., & Abrams, R. (2007). Priming of semantic classifications by novel subliminal prime words. *Consciousness and Cognition*, 16, 63–83.
- Klauer, K. C., Mierke, J., & Musch, J. (2003). The positivity proportion effect: A list context effect in masked affective priming. *Memory & Cognition*, 31, 953–967.
- Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–50). Mahwah, NJ: Erlbaum.
- Klauer, K. C., Musch, J., & Eder, A. (2005). Priming of semantic classification: Late and response-related, or earlier and more central? *Psychonomic Bulletin & Review*, 12, 897–903.
- Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (2009). Contrast effects in spontaneous evaluations: A psychophysical account. *Journal of Personality and Social Psychology*, 96, 265–287.
- Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition*, 88, 223–242.
- Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *Journal of Experimental Psychology: General*, 133, 475–493.
- Lleras, A., & Enns, J. T. (2005). Updating a cautionary tale of masked priming: Reply to Klapp (2005). *Journal of Experimental Psychology: General*, 134, 436–440.
- 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, 495–500.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330.
- Miller, G. A. (1964). *Mathematics and psychology*. New York, NY: Wiley.
- Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 861–872.
- Praamstra, P., & Seiss, E. (2005). The neurophysiology of response competition: Motor cortex activation and inhibition following subliminal response priming. *Journal of Cognitive Neuroscience*, 17, 483–493.
- Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, 132, 416–442.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Reynvoet, B., Gevers, W., & Caessens, B. (2005). Unconscious primes activate motor codes through semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 991–1000.
- Rothermund, K. (2003). Motivation and attention: Incongruent effects of feedback on the processing of valence. *Emotion*, 3, 223–238.
- Schlaghecken, F., Blagrove, E., & Maylor, E. A. (2008). No difference between conscious and nonconscious visuomotor control: Evidence from perceptual learning in the masked prime task. *Consciousness and Cognition*, 17, 84–93.
- Schlaghecken, F., & Eimer, M. (2000). A central-peripheral asymmetry in masked priming. *Perception & Psychophysics*, 62, 1367–1382.
- Schlaghecken, F., & Eimer, M. (2002). Motor activation with and without inhibition: Evidence for a threshold mechanism in motor control. *Perception & Psychophysics*, 64, 148–162.

- Schlaghecken, F., & Eimer, M. (2006). Active masks and active inhibition: A comment on Lleras and Enns (2004) and on Verleger, Jaśkowski, Aydemir, van der Lubbe, and Groen (2004). *Journal of Experimental Psychology: General*, *135*, 484–494.
- Schlaghecken, F., Klapp, S., & 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*, 515–521.
- Sumner, P. (2007). Negative and positive masked-priming: Implications for motor inhibition. *Advances in Cognitive Psychology*, *3*, 317–326.
- Sumner, P., & Brandwood, T. (2008). Oscillations in motor priming: Positive rebound follows inhibitory phase in the masked prime paradigm. *Journal of Motor Behavior*, *40*, 484–489.
- Sumner, P., & Husain, M. (2008). At the edge of consciousness: Automatic motor activation and voluntary control. *Neuroscientist*, *14*, 474–486.
- Sumner, P., Tsai, P.-C., Yu, K., & Nachev, P. (2006). Attentional modulation of sensorimotor processes in the absence of perceptual awareness. *Proceedings of the National Academy of Sciences*, *103*, 10520–10525.
- Townsend, J. T., & Ashby, F. G. (1993). *The stochastic modeling of elementary psychological processes*. New York, NY: Cambridge University Press.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.
- Verleger, R., Jaśkowski, P., Aydemir, A., van der Lubbe, R. H. J., & Groen, M. (2004). Qualitative differences between conscious and non-conscious processing? On inverse priming induced by masked arrows. *Journal of Experimental Psychology: General*, *133*, 494–515.
- Voss, A., Leonhart, R., & Stahl, C. (2007). How to make your own response boxes: A step-by-step guide for the construction of reliable and inexpensive parallel-port response pads from computer mice. *Behaviour Research Methods*, *39*, 797–801.
- Wittenbrink, B. (2007). Measuring attitudes through priming. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies* (pp. 17–58). New York, NY: Guilford Press.

Appendix

Analyses of Correct Response Latencies and Percentage-Correct Scores

Tables in the Appendix show mean latencies and mean percentage-correct scores. The participants' percentage-correct scores were arcsine-transformed for the statistical analyses, and we present both mean arcsine-transformed values (for comparison with the effects reported in the statistical analyses) as well as the mean values retransformed to the original percentage-correct scale to facilitate the assessment of accuracy levels.

Experiment 1

Mean correct response latencies and percentage-correct scores are shown in Table A1 as a function of prime, Stimulus 2, and target. An analysis of variance of the correct response latencies with these factors showed a significant main effect of Stimulus 2, $F(3, 57) = 466.46$, $p < .01$, $\epsilon = .91$, $\eta_p^2 = .96$. Overall, responses were faster in trials with response-valent compatible Stimulus 2 than in trials with response-valent incompatible Stimulus 2, with response latencies for trials with relevant masks and irrelevant masks between. There was also an interaction of prime and Stimulus 2, $F(3, 57) = 3.36$, $p = .03$, $\epsilon = .84$, $\eta_p^2 = .15$, reflecting faster responses for right-pointing arrows as prime relative to left-pointing arrows in trials with the relevant mask, with no effect of prime in trials with other kinds of Stimulus 2. Finally, there is a significant three-way interaction of all factors that is numerically and statistically the same as the main effect of Stimulus 2 on compatibility effects exerted by the prime on the target, as reported in the body of the article.

An analysis of variance of percentage-correct scores (arcsine-transformed) with the same factors showed a significant main effect of Stimulus 2, $F(3, 57) = 84.21$, $p < .01$, $\epsilon = .67$, $\eta_p^2 = .82$. Overall, responses were least accurate following the response-valent incompatible Stimulus 2 and most accurate following the response-valent compatible Stimulus 2 with accuracies for trials

with relevant masks and irrelevant masks between. There was also a small, but significant main effect of target, $F(1, 19) = 4.79$, $p = .04$, $\eta_p^2 = .20$, so that responses to left-pointing arrows were somewhat more accurate than responses to right-pointing arrows. Finally, there is a significant three-way interaction of all factors that is numerically and statistically the same as the main effect of Stimulus 2 on compatibility effects exerted by the prime on the target as reported in the body of the article.

Experiment 2

Mean correct response latencies and percentage-correct scores are shown in Table A2 as a function of group, prime, Stimulus 2, and target. An analysis of variance of the correct response latencies with repeated measures on all of these factors other than the group factor was computed. Group was entered as a between-participants factor. There was a significant main effect of Stimulus 2, $F(3, 171) = 565.89$, $p < .01$, $\epsilon = .73$, $\eta_p^2 = .91$: Response latencies were largest for trials with response-valent incompatible Stimulus 2 and smaller for trials with response-valent compatible Stimulus 2, with response latencies for trials with relevant and irrelevant masks falling between. A main effect of group, $F(1, 57) = 7.59$, $p < .01$, $\eta_p^2 = .12$, showed that the slow group responded significantly faster than the fast group. An interaction of Stimulus 2 and group, $F(3, 171) = 11.80$, $p < .01$, $\epsilon = .73$, $\eta_p^2 = .17$, went back to the fact that the effects of Stimulus 2 were more pronounced in the fast group than in the slow group. A main effect of target, $F(1, 57) = 8.62$, $p < .01$, $\eta_p^2 = .13$, reflected faster responses to right double arrows than to left double arrows. Effects involving the interaction of prime and target are numerically and statistically the same as corresponding effects on compatibility effects reported in the body of the article.

(Appendix continues)

Table A1
Experiment 1

Prime	Target	Stimulus 2			
		Relevant	Irrelevant	Incompatible	Compatible
Latency					
Right	Right	404	392	438	299
Right	Left	380	383	424	350
Left	Right	379	377	422	351
Left	Left	418	394	442	302
Accuracy data (arcsine scale)					
Right	Right	1,372	1,415	1,238	1,535
Right	Left	1,514	1,472	1,165	1,520
Left	Right	1,479	1,441	1,129	1,520
Left	Left	1,389	1,411	1,287	1,563
Accuracy data (percentage-correct scale)					
Right	Right	96.10	97.58	89.33	99.87
Right	Left	99.68	99.03	84.41	99.75
Left	Right	99.16	98.33	81.73	99.75
Left	Left	96.72	97.48	92.17	99.99

Note. Mean correct response latencies (ms) and mean percentage-correct scores (arcsine-transformed multiplied by 1,000 and retransformed to the percentage-correct scale) as a function of prime, Stimulus 2, and target.

An analysis of variance of percentage-correct scores (arcsine-transformed) with the same factors showed a significant main effect of prime, $F(1, 57) = 4.20, p = .05, \eta_p^2 = .07$, with more accurate responses in trials with right arrows as prime. There was also a main effect of Stimulus 2, $F(3, 171) = 87.70, p < .01, \epsilon =$

$.63, \eta_p^2 = .61$, with most accurate responses following the response-valent compatible Stimuli 2 and the least accurate response following response-valent incompatible Stimuli 2, with accuracies for trials with relevant and irrelevant masks between. This effect was moderated by group, $F(3, 171) = 7.80, p < .01$,

Table A2
Experiment 2

Prime	Target	Fast group—Stimulus 2				Slow group—Stimulus 2			
		Relevant	Irrelevant	Incompatible	Compatible	Relevant	Irrelevant	Incompatible	Compatible
Latency									
Right	Right	397	385	403	293	369	360	389	310
Right	Left	369	366	409	348	352	335	375	310
Left	Right	359	361	401	342	345	336	374	310
Left	Left	406	399	419	292	372	361	391	314
Accuracy data (arcsine scale)									
Right	Right	1,422	1,429	1,318	1,562	1,429	1,476	1,357	1,544
Right	Left	1,486	1,526	1,202	1,509	1,503	1,522	1,431	1,544
Left	Right	1,474	1,489	1,268	1,523	1,472	1,502	1,433	1,547
Left	Left	1,371	1,406	1,337	1,536	1,389	1,420	1,362	1,561
Accuracy data (percentage-correct scale)									
Right	Right	97.81	98.00	93.74	99.99	98.00	99.10	95.52	99.93
Right	Left	99.29	99.79	87.02	99.62	99.53	99.76	98.05	99.93
Left	Right	99.06	99.33	91.12	99.77	99.02	99.53	98.11	99.94
Left	Left	96.05	97.31	94.63	99.88	96.74	97.74	95.71	99.99

Note. Mean correct response latencies (ms) and mean percentage-correct scores (arcsine-transformed multiplied by 1,000 and retransformed to the percentage-correct scale) as a function of group, prime, Stimulus 2, and target.

(Appendix continues)

$\epsilon = .85$, $\eta_p^2 = .12$, the Stimulus 2 effects being more pronounced in the fast group. A main effect of group, $F(1, 57) = 4.56$, $p = .04$, $\eta_p^2 = .07$, indicates that there were more errors in the fast group than in the slow group. An interaction of prime and Stimulus 2, $F(3, 171) = 5.49$, $p < .01$, $\epsilon = .90$, $\eta_p^2 = .09$, showed that the accuracy advantage for the right arrow as prime was restricted to trials with mask as Stimulus 2. Effects involving the interaction of prime and target are numerically and statistically the same as corresponding effects on compatibility effects reported in the body of the article.

Experiment 3

Mean correct response latencies and percentage-correct scores are shown in Table A3 as a function of group, prime category,

Stimulus 2 category, target category, and stimulus family. An analysis of variance of the correct response latencies with repeated measures on all of these factors other than the group factor was computed. In the Appendix, we report effects and interactions not involving the interaction of prime and target because these effects correspond to effects on compatibility effects already discussed in the body of the text. We report all other significant effects and interactions but do not attempt to interpret interactions of higher order interactions than three-way interactions.

Trials with letter targets were responded to slower than were trials with arrow targets, a main effect of stimulus family, $F(1, 37) = 187.57$, $p < .01$, $\eta_p^2 = .84$. The effect interacted with group, $F(1, 37) = 17.93$, $p < .01$, $\eta_p^2 = .33$, so that the difference between letters and arrows was more pronounced for the mixed-family

Table A3
Experiment 3

Prime	Target	Single-family group—Stimulus 2				Mixed-family group—Stimulus 2			
		Relevant	Irrelevant	Incompatible	Compatible	Relevant	Irrelevant	Incompatible	Compatible
Latency									
Arrow stimulus family									
Right	Right	389	379	407	305	382	366	392	363
Right	Left	348	344	381	315	361	351	372	345
Left	Right	350	350	379	313	356	347	376	354
Left	Left	382	376	395	314	373	367	400	365
Letter stimulus family									
Right	Right	419	377	430	311	394	405	430	406
Right	Left	367	404	416	342	397	414	427	427
Left	Right	389	396	412	322	396	394	409	389
Left	Left	377	395	440	314	416	422	440	430
Accuracy data (arcsine scale)									
Arrow stimulus family									
Right	Right	1,409	1,496	1,402	1,570	1,447	1,478	1,448	1,473
Right	Left	1,528	1,535	1,400	1,542	1,500	1,548	1,516	1,538
Left	Right	1,524	1,535	1,431	1,570	1,528	1,518	1,496	1,492
Left	Left	1,391	1,426	1,360	1,529	1,507	1,506	1,428	1,434
Letter stimulus family									
Right	Right	1,208	1,384	1,283	1,487	1,410	1,358	1,270	1,383
Right	Left	1,487	1,404	1,482	1,550	1,423	1,342	1,325	1,311
Left	Right	1,386	1,451	1,336	1,545	1,417	1,457	1,350	1,436
Left	Left	1,385	1,345	1,212	1,539	1,264	1,280	1,310	1,286
Accuracy data (percentage-correct scale)									
Arrow stimulus family									
Right	Right	97.41	99.43	97.16	100.00	98.48	99.14	98.50	99.04
Right	Left	99.82	99.87	97.12	99.92	99.49	99.95	99.70	99.89
Left	Right	99.78	99.87	98.05	100.00	99.82	99.72	99.44	99.38
Left	Left	96.82	97.91	95.61	99.83	99.59	99.58	97.97	98.14
Letter stimulus family									
Right	Right	87.38	96.56	91.92	99.30	97.43	95.53	91.22	96.51
Right	Left	99.30	97.25	99.21	99.95	97.84	94.88	94.08	93.39
Left	Right	96.61	98.57	94.57	99.93	97.64	98.72	95.21	98.19
Left	Left	96.58	94.98	87.70	99.90	90.88	91.76	93.37	92.09

Note. Mean correct response latencies (ms) and mean percentage-correct scores (arcsine-transformed multiplied by 1,000 and retransformed to the percentage-correct scale) as a function of group, prime category, Stimulus 2 category, target category, and stimulus family.

(Appendix continues)

group than for the single-family group. There was also a main effect of Stimulus 2 category, $F(3, 111) = 178.41, p < .01, \epsilon = .67, \eta_p^2 = .83$: Responses were fastest following response-valent compatible Stimuli 2 and slowest following response-valent incompatible Stimuli with response latencies in trials with relevant and irrelevant masks between. The effect was moderated by group, $F(3, 111) = 89.46, p < .01, \epsilon = .67, \eta_p^2 = .71$, indicating that the single-family group profited from response-valent compatible Stimuli 2 much more than the mixed-family group. Responses to right targets were faster than were responses to left targets, $F(1, 37) = 4.32, p = .05, \eta_p^2 = .11$, a main effect of target that interacted with group, $F(1, 37) = 6.87, p = .01, \eta_p^2 = .16$: The difference was apparent only in the mixed-family group. Stimulus 2 category interacted with stimulus family, $F(3, 111) = 5.76, p < .01, \epsilon = .82, \eta_p^2 = .14$: Responses in letter-target trials with relevant masks were faster than were responses in letter-target trials with irrelevant mask and vice versa in trials with arrow targets, and the effect of stimulus family was reduced in trials with response-valent compatible Stimuli 2. This two-way interaction itself interacted with group, $F(3, 111) = 11.10, p < .01, \epsilon = .82, \eta_p^2 = .23$: The effects of Stimulus 2 on the differences between the two stimulus families were reduced in the mixed-family group. An interaction of stimulus family and target category, $F(1, 37) = 5.27, p = .03, \eta_p^2 = .13$, revealed that the latency advantage for right targets was restricted to letter targets. An interaction of Stimulus 2 category and target, $F(3, 111) = 10.44, p < .01, \epsilon = .96, \eta_p^2 = .22$, went back to the fact that left targets were responded to faster than were right targets, following relevant masks, but slower for all other kinds of Stimuli 2. The interaction was moderated by group, $F(3, 111) = 2.95, p = .04, \epsilon = .96, \eta_p^2 = .02$, so that the above two-way interaction was apparent in the single-family group but not in the mixed-family group. An interaction of prime, Stimulus 2 category, and stimulus family, $F(3, 111) = 2.84, p = .04, \epsilon = .98, \eta_p^2 = .07$, indicated that left primes induced faster overall responses than did right primes for arrow targets preceded by relevant masks and for letter targets preceded by response-compatible Stimuli 2, whereas right primes induced faster overall responses than did left primes for arrow targets preceded by response-compatible Stimuli 2, with little effect of prime in other conditions. There was also a complex four-way interaction of prime, Stimulus 2 category, stimulus family, and group, $F(3, 111) = 4.52, p = .01, \epsilon = .98, \eta_p^2 = .11$. An interaction of Stimulus 2 category and target category, $F(3, 111) = 10.44, p < .01, \epsilon = .96, \eta_p^2 = .22$, revealed that the latency advantage for right targets over left targets reversed for relevant masks. This reversal was apparent in both groups, but there was little effect of target for other kinds of Stimulus 2 in the single-family group, causing a three-way interaction of Stimulus 2 category, target category, and group, $F(3, 111) = 2.95, p = .04, \epsilon = .96, \eta_p^2 = .07$. The interaction of Stimulus 2 category and target category was also moderated by stimulus family, $F(3, 111) = 7.17, p < .01, \epsilon = .77, \eta_p^2 = .16$: There was little effect of target category for arrow targets, whereas the two-way interaction was apparent only for letter targets. There was also a complex four-way interaction

involving Stimulus 2 category, target category, stimulus family, and group, $F(3, 111) = 5.37, p < .01, \epsilon = .77, \eta_p^2 = .13$.

Considering percentage-correct scores (arcsine-transformed), trials with letter targets were responded to less accurately than were trials with arrow targets, $F(1, 37) = 122.04, p < .01, \eta_p^2 = .77$, a main effect of stimulus family that interacted with group, $F(1, 37) = 12.08, p < .01, \eta_p^2 = .25$: The effect of stimulus family was more pronounced in the mixed-family group than in the single-family group. Furthermore, accuracy was highest following response-valent compatible Stimuli 2 and lowest following response-valent incompatible Stimuli 2, with accuracies in trials following relevant and irrelevant mask in between, $F(3, 111) = 17.86, p < .01, \epsilon = .58, \eta_p^2 = .33$. Stimulus 2 category interacted with group, $F(3, 111) = 12.36, p < .01, \epsilon = .58, \eta_p^2 = .25$, the effects of Stimulus 2 compatibility for response-valent Stimuli 2 being much more pronounced in the single-family group than in the mixed-family group. Target category also interacted with group, $F(1, 37) = 4.55, p = .04, \eta_p^2 = .11$, with more accurate responses to right targets than left targets only in the mixed-family group. For letter targets, there was also a left over right advantage in the single-family group, and the target Category \times Group interaction was reversed for arrow targets, leading to a three-way interaction of target category, group, and stimulus family, $F(1, 37) = 15.06, p < .01, \eta_p^2 = .29$. Finally, there was a three-way interaction of Stimulus 2 category, stimulus family, and target category, $F(3, 111) = 2.99, p = .04, \epsilon = .91, \eta_p^2 = .08$: For letter targets, right targets were responded to more accurately than were left targets following an irrelevant mask, and vice versa following a relevant mask; in contrast for arrow targets, there was little evidence for an interaction of Stimulus 2 category and target category.

Experiment 4

In Table A4, mean correct response latencies and percentage-correct scores are shown as a function of group, Stimulus 2 category, prime category, target category, and stimulus family. An analysis of variance of the correct response latencies with repeated measures on all of these factors other than the group factor was computed. Responses in the variable-mapping group were slower than responses in the fixed-mapping group, $F(1, 27) = 70.70, p < .01, \eta_p^2 = .72$, a main effect that was moderated by an interaction with stimulus family, $F(1, 27) = 74.59, p < .01, \eta_p^2 = .73$. Responses were faster for arrow targets than for letter targets in the fixed-mapping group, but the reverse was true in the variable-mapping group. This two-way interaction was itself moderated by target category, $F(3, 81) = 5.93, p = .02, \eta_p^2 = .18$, the two-way interaction being more pronounced for left targets than for right targets. There was also a complex four-way interaction involving these three factors and Stimulus 2 category, $F(3, 81) = 3.62, p = .03, \epsilon = .75, \eta_p^2 = .12$, which we do not attempt to interpret. There was a main effect of Stimulus 2 category, $F(3, 81) = 23.32, p < .01, \epsilon = .71, \eta_p^2 = .46$, with responses slowest in trials with response-valent incompatible Stimuli 2, followed by trials with response-valent compatible Stimuli 2, with responses fastest in

Table A4
Experiment 4

Prime	Target	Fixed-map group—Stimulus 2				Variable-mapping group—Stimulus 2			
		Relevant	Irrelevant	Incompatible	Compatible	Relevant	Irrelevant	Incompatible	Compatible
Latency									
Arrow stimulus family									
Right	Right	393	375	392	376	691	696	698	709
Right	Left	365	364	390	358	684	682	681	682
Left	Right	377	362	394	364	662	676	682	694
Left	Left	375	372	399	380	699	678	694	710
Letter stimulus family									
Right	Right	420	418	458	422	657	658	668	661
Right	Left	432	434	451	443	641	631	644	647
Left	Right	405	413	449	422	645	639	659	639
Left	Left	433	443	466	449	636	643	661	660
Accuracy data (arcsine scale)									
Arrow stimulus family									
Right	Right	1,541	1,520	1,513	1,484	1,080	1,105	1,096	1,057
Right	Left	1,540	1,525	1,457	1,505	1,105	1,125	1,160	1,035
Left	Right	1,504	1,520	1,505	1,490	1,120	1,156	1,157	1,088
Left	Left	1,500	1,490	1,448	1,520	1,089	1,083	1,087	1,050
Letter stimulus family									
Right	Right	1,324	1,343	1,263	1,405	1,243	1,245	1,214	1,265
Right	Left	1,374	1,301	1,352	1,374	1,256	1,306	1,240	1,221
Left	Right	1,438	1,348	1,280	1,377	1,253	1,219	1,232	1,272
Left	Left	1,235	1,203	1,378	1,278	1,282	1,224	1,213	1,255
Accuracy data (percentage-correct scale)									
Arrow stimulus family									
Right	Right	99.91	99.74	99.67	99.24	77.81	79.84	79.12	75.83
Right	Left	99.91	99.79	98.71	99.57	79.84	81.41	84.02	73.91
Left	Right	99.55	99.74	99.57	99.35	81.01	83.75	83.84	78.47
Left	Left	99.50	99.35	98.49	99.74	78.53	78.01	78.38	75.21
Letter stimulus family									
Right	Right	94.01	94.92	90.83	97.28	89.61	89.78	87.81	90.93
Right	Left	96.16	92.91	95.30	96.16	90.43	93.14	89.46	88.23
Left	Right	98.25	95.10	91.76	96.28	90.26	88.10	88.95	91.32
Left	Left	89.12	87.10	96.32	91.69	91.90	88.44	87.71	90.37

Note. Mean correct response latencies (ms) and mean percentage-correct scores (arcsine-transformed multiplied by 1,000 and retransformed to the percentage-correct scale) as a function of group, prime category, Stimulus 2 category, target category, and stimulus family.

trials with masks. The main effect was moderated by group, $F(3, 81) = 8.24, p < .01, \varepsilon = .71, \eta_p^2 = .23$, the effects of Stimulus 2 being more pronounced in the fixed-mapping group than in the variable-mapping group. Effects involving the interaction of prime and target are numerically and statistically the same as corresponding effects on compatibility effects as reported in the body of the article.

The analogous analysis of variance of the percentage-correct scores (arcsine-transformed) revealed a main effect of group, $F(1, 27) = 17.11, p < .01, \eta_p^2 = .39$, indicating that more errors were made in the variable-mapping group than in the fixed-mapping group. The interaction of group and stimulus family, $F(1, 27) = 64.78, p < .01, \eta_p^2 = .71$, was significant as well: Whereas responses were more accurate for arrow targets than for letter targets in the fixed-mapping group, the reverse pattern occurred in

the variable-mapping group. There was also a complex four-way interaction of stimulus family, group, Stimulus 2 category, and group, $F(3, 81) = 3.60, p = .02, \varepsilon = .93, \eta_p^2 = .12$, which we do not attempt to interpret. Effects involving the interaction of prime and target are numerically and statistically the same as corresponding effects on compatibility effects as reported in the body of the article.

Experiment 5

In Table A5, mean correct response latencies and percentage-correct scores are shown as a function of prime category, target category, and SOA. An analysis of variance of correct response latencies with repeated measures on all of these factors revealed a significant effect of SOA, $F(2, 190) = 206.09, p < .01, \varepsilon = .81$,

(Appendix continues)

$\eta_p^2 = .68$, reflecting increasingly faster responses, the longer the SOA. The significant three-way interaction of all three factors is numerically and statistically the same as the effect of SOA on compatibility effects reported in the body of the article.

The analogous analysis of variance of percentage-correct scores (arcsine-transformed) also revealed a main effect of SOA, $F(2, 190) = 6.30, p < .01, \epsilon = .93, \eta_p^2 = .06$, reflecting an increase in errors as SOA increased. There was a main effect of target category, $F(1, 95) = 7.18, p < .01, \eta_p^2 = .07$, with more accurate responses for the targets requiring the up rather than down response. The significant three-way interaction of all three factors is numerically and statistically the same as the effect of SOA on compatibility effects reported in the body of the article.

Experiment 6

Mean correct response latencies and percentage-correct scores are shown in Table A6 as a function of prime category, target category, and stream length. An analysis of variance of the correct response latencies with all of these factors was conducted. A significant main effect of prime, $F(1, 100) = 7.54, p < .01, \eta_p^2 = .07$, reflected faster

Table A5
Experiment 5

Prime	Target	Stimulus-onset asynchrony (ms)		
		120	240	360
Latency				
Up	Up	392	378	365
Up	Down	401	374	368
Down	Up	400	371	360
Down	Down	392	379	367
Accuracy data (arcsine scale)				
Up	Up	1,401	1,370	1,376
Up	Down	1,366	1,357	1,348
Down	Up	1,361	1,371	1,367
Down	Down	1,387	1,362	1,329
Accuracy data (percentage-correct scale)				
Up	Up	97.13	96.02	96.24
Up	Down	95.86	95.52	95.10
Down	Up	95.70	96.07	95.89
Down	Down	96.65	95.70	94.28

Note. Mean correct response latencies (ms) and mean percentage-correct scores (arcsine-transformed multiplied by 1,000 and retransformed to the percentage-correct scale) as a function of prime category, target category, and stimulus-onset asynchrony.

Table A6
Experiment 6

Prime	Target	Stream length			
		2-3	4-5	6-7	8-9
Latency					
Up	Up	674	661	657	659
Up	Down	669	665	659	658
Down	Up	673	668	664	666
Down	Down	677	661	657	654
Accuracy data (arcsine scale)					
Up	Up	1,329	1,317	1,319	1,333
Up	Down	1,330	1,321	1,323	1,351
Down	Up	1,340	1,329	1,302	1,312
Down	Down	1,320	1,329	1,342	1,339
Accuracy data (percentage-correct scale)					
Up	Up	94.27	93.70	93.80	94.45
Up	Down	94.31	93.89	94.01	95.25
Down	Up	94.77	94.28	92.96	93.45
Down	Down	93.85	94.27	94.87	94.70

Note. Mean correct response latencies (ms) and mean percentage-correct scores (arcsine-transformed multiplied by 1,000 and retransformed to the percentage-correct scale) as a function of prime category, target category, and stream length.

responses in trials with primes pointing upward than in trials with primes pointing downward. The main effect of stream length was significant as well, $F(3, 300) = 49.67, p < .01, \epsilon = .78, \eta_p^2 = .33$, reflecting increasingly shorter reaction times, the longer the stream. An interaction between target category and stream length, $F(3, 300) = 2.69, p = .05, \epsilon = .96, \eta_p^2 = .03$, went back to the fact that downward-pointing targets were responded to faster for the longest stream, whereas there was little differences between up and down targets for shorter streams. Effects involving the interaction of prime and target are numerically and statistically the same as the corresponding compatibility effects reported in the body of the article.

An analysis of variance of the percentage-correct scores (arcsine-transformed) with the same factors showed a significant interaction of target and stream length, $F(3, 300) = 3.44, p = .02, \epsilon = .93, \eta_p^2 = .03$: Fewer errors were made for targets pointing downward than for targets pointing upward for longer streams, whereas the reverse was true for the shortest streams.

Received October 26, 2009

Revision received March 26, 2010

Accepted March 26, 2010 ■