Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming

Sachiko Kinoshita

Macquarie University, Australia

Michael C. Mozer

University of Colorado, Boulder

and

Kenneth I. Forster

University of Arizona, Tucson

Running head: Dynamic adaptation to trial difficulty

As at April 26, 2011

Contact Information:

Sachiko Kinoshita
Macquarie Centre for Cognitive Science and
Department of Psychology
Macquarie University
Sydney, NSW 2109
Australia
Email: sachiko.kinoshita@mq.edu.au
Fax: 61 2 9850 6059

Abstract

In reaction time research, there has been an increasing appreciation that response initiation processes are sensitive to recent experience, and in particular, the difficulty of previous trials. From this perspective, we propose an explanation for a perplexing property of masked priming: although primes are not consciously identified, facilitation of target processing by a related prime is magnified in a block containing a high proportion of related primes relative to a block containing the opposite mix (Bodner & Masson, 2001). In the present study, we explore this phenomenon with a parity (even/odd) decision task in which a prime (e.g., two) precedes a target that can be either congruent (e.g., four) or incongruent (e.g., three). We show that the effect of congruence proportion with masked primes cannot be explained in terms of the blockwise prime-target contingency. Specifically, with masked primes there is no congruency disadvantage in a block containing a high proportion of incongruent primes, but a congruency advantage when the block contains an equal proportion of congruent and incongruent primes. In qualitative contrast, visible primes are sensitive to the blockwise prime-target contingency. We explain the relatedness proportion effect found with masked primes in terms of a model according to which response initiation processes adapt to the statistical structure of the environment, specifically the difficulty of recent trials. We support our account with an analysis at the level of individual trials using the linear mixed effects model. (235 words)

Keywords: masked priming, prime validity, proportion effect, parity decision, memory recruitment account; contingency learning

For over half a century, sequential dependencies have been described in experimental paradigms requiring individuals to perform a task repeatedly or to perform a series of tasks (e.g., Anderson, 1960; Estes, 1950; Laming, 1979; Remington, 1969). Sequential dependencies involve an incidental influence of recent stimuli and responses on the current trial. Diverse behavioral measures are affected, including response latency, accuracy, type of errors produced, and interpretation of ambiguous stimuli. Sequential dependencies are ubiquitous and occur across a variety of experimental tasks, including speeded response, judgment, and identification, and across all components of the cognitive architecture, including perception (Maloney, Dal Martello, Sahm, & Spillmann, 2005), selective attention (Kristjansson, 2006), categorization (Jones & Sieck, 2003), language (Bock & Griffin, 2000), decision making (Jesteadt, Luce, & Green, 1977), action selection (Dixon & Glover. 2009), and executive control (Gratton, Coles, & Donchin 1992).

Although some sequential effects can be considered in terms of stimulus- or response-specific priming, many forms involve abstractions of recent experience. For example, in speeded word or object naming, the difficulty of recent items—whether due to frequency, orthographic regularity, or number of syllables—influence latency to naming: an individual slows down following a hard item and speeds up following an easy item (Lupker, Brown & Colombo, 1997; Meyer, Roelofs, & Levelt, 2003; Taylor & Lupker, 2001). This effect occurs even in the absence of errors and is thus different from posterror slowing (Rabbitt, 1989).

Recent theoretical perspectives characterize sequential dependencies in terms of an individual's adaptation to the statistical structure of a nonstationary environment

(Jones & Sieck, 2003; Mozer, Kinoshita, & Shettel, 2007; Wilder, Jones, & Mozer, 2010; Yu & Cohen, 2009). Nonstationarity implies that recent events provide a stronger predictor of the future than events further back in time. Sequential dependencies suggest that what might appear to be intrinsic uncertainty in human behavior can instead be understood in terms of a systematic, adaptive response to a changing world.

Sequential dependencies also offer an opportunity to gain a more fundamental, mechanistic understanding of empirical phenomena that are cast in terms of blockwise manipulations. For example, in word naming, response initiation slows down for easy words when they are presented in a mixed block containing both easy and hard words relative to a pure block of easy words; likewise, response initiation speeds up for hard words when they are presented in a mixed block relative to a pure block of hard words (e.g., Lupker et al., 1997). Although this blocking effect can be described in terms of the composition of a block, the effect can be more fully explained in terms of a trial-to-trial adaptation to item difficulty (Mozer, Kinoshita, & Davis, 2004; Jones, Mozer, & Kinoshita, 2009). The trial-to-trial perspective facilitates mechanistic accounts that are not readily apparent when phenomena are described in terms of whole-block statistics.

In this article, our goal is to provide a simple mechanistic explanation for a phenomenon that seems somewhat perplexing when considered from the perspective of a blocking manipulation: the effect of block composition on the effectiveness of masked primes.

Relatedness proportion effect in masked priming. In the masked priming procedure, a prime is presented very briefly, and preceded by a forward-mask (typically #

signs) and followed (and backward-masked) by a target. Despite the fact the prime is not consciously perceived, primes that are *related* to the target in various ways (e.g., identity/repetition – *table-TABLE*, orthographic – *bontrast-CONTRAST*, semantic – *eagle-HAWK*) facilitate responses to the target relative to *unrelated* primes. Because the prime is not available for conscious report, the masked priming procedure is a popular tool used by researchers interested in unconscious, automatic processes (e.g., Forster & Davis, 1984; Dehaene, Naccache, Le Clec, et al., 1998; Kouider & Dehaene, 2007).

Although the finding of masked priming effect is now well-accepted, the finding that the proportion of related primes affects the size of priming effect is more controversial, and surprising. Bodner and Masson (2001) have reported that the size of masked priming effect is reduced in a block containing a high proportion (.8) of unrelated prime trials and low proportion (.2) of related prime trials relative to a block containing an opposite mix. Bodner and colleagues replicated this effect of relatedness proportion on masked priming in several speeded response tasks, including lexical decision (Bodner & Masson, 2001, 2003)[1], naming (Bodner & Masson, 2004), and number magnitude judgment and number parity (even/odd) judgment (Bodner & Dypvik, 2005). Bodner and colleagues termed the effect the *prime validity effect*, but we use the theoretically neutral term *the relatedness proportion effect* here. Similar proportion effects with masked primes have also been reported with other speeded response tasks involving simple perceptual decisions, for example, with arrow stimuli (“<<” or “>>”) for which subjects indicate which way they are pointing (Klapp, 2007), and with square stimuli for which

[1] Note however that the finding of proportion effects in the lexical decision task is mixed (refer to Table 1 in Bodner & Masson, 2001). We will return to this in the General Discussion.

subjects decide whether a line segment is missing (Jaskowski, Skalska, & Verleger, 2003).

As Bodner and Masson (2001) noted, the effect of relatedness proportion on masked priming is counterintuitive. Relatedness proportion effects are well-documented in the semantic priming literature (see Neely, 1991 for a review), but in these studies the prime was clearly visible. The relatedness proportion effects with visible primes are explained in terms of controlled, strategic processes, where the prime is used to intentionally generate expectancies about the target. In masked priming, people are typically unaware of the presence of the prime, let alone its identity. How could they adapt to the proportion of congruent trials in the absence of the awareness of prime identity? The present study investigates this question, using the parity judgment task used by Bodner and Dypvik (2005). In this task, subjects decide whether a number target (which may be a digit, e.g., 2, or a spelled-out word, e.g., THREE) is odd or even. The congruence effect refers to the faster decision to the target when the prime shares the parity with the target (e,g, one-THREE) relative to when they do not (e.g., two-THREE). The congruency proportion effect (what Bodner and Dypvik, 2005, called the prime validity effect) is the finding that the size of congruence effect is reduced in the block containing a low proportion of congruent trials relative to a block containing a high proportion of congruent trials.

The paper is organized into two parts. The first part considers the possibility that the congruency proportion effect reflects adaptation to blockwise prime utility. By blockwise prime utility, we mean the predictability of the target *response* from the prime based on the blockwise statistics of prime-target contingency. We first point out that congruency proportion and prime-target contingency are not the same. We then show that the priming effect with visible primes tracks prime-target contingency. These results observed with visible primes are however qualitatively different from the pattern of congruency proportion effects observed with masked primes, and hence challenge the interpretation that masked priming is sensitive to the utility of the prime information across trials. We take this dissociation to argue against the possibility that people can adapt to the blockwise utility of the prime when the prime is not consciously identifiable. In the second part of the paper, we describe an alternative account of the congruency proportion effect based on our model of optimal response initiation called the Adaptation-to-the-Statistics of the Environment (ASE) model (Jones, Mozer, & Kinoshita, 2009; Kinoshita, Forster, & Mozer, 2008; Mozer, Kinoshita, & Davis, 2004). We report analyses using a linear mixed effects model, examining data at the level of individual trials to test our account. These analyses support the ASE account in revealing a large effect of previous trial difficulty, and how it is modulated by the difficulty of the item (congruent vs. incongruent trials), and by the difficulty of task environment (the proportion of congruent trials).

Congruency proportion vs. prime utility. The relatedness proportion effect observed with masked primes is widely believed to reflect the adaptation of the cognitive system to the utility of prime information across trials (e.g., Balota, Yap, Cortese, & Watson, 2008; Schmidt, Crump, Cheesman & Besner, 2007). The term “the prime validity effect” implies this interpretation, and the high-congruent-proportion block is usually called the “high-validity” block, and to the low-proportion-congruent block the "low validity" block. However, this interpretation may be questioned. The low-proportion-congruent block contains 20% congruent trials and 80% incongruent trials. In the parity (odd-even) decision task, this means that the primes actually predict the (opposite) response to the target: if the prime is odd, the target would be even in 80% of the cases, and vice versa. That is, the low-proportion-congruent block may be considered to have *high* prime utility. If subjects utilize the prime episode to facilitate responding to the target, what should be observed in this block is a *reverse congruence effect*, i.e., performance should be better (faster/more accurate) in the incongruent trials than the congruent trials. What Bodner and Dypvik (2005) found, on the contrary, was a *positive* congruence effect, but reduced relative to a high-proportion congruent block.

Prime utility is in fact the lowest when the prime does not predict either response to the target, i.e., when there is an equal proportion of congruent trials and incongruent trials. In this case, subjects should ignore the prime parity, as it has no bearing on the response to the target. Consequently, there should be no congruence effect in this case. In contradiction of this expectation, robust congruence effects in the parity decision task have been observed in previous studies using an equal proportion of congruent and incongruent trials (e.g., Reynvoet, et al., 2002; Fabre & Lemaire, 2005), as is typical in masked priming experiments. In sum, proportion effects with masked primes in the parity task do not resemble the pattern expected from the view that subjects were adapting to blockwise prime-target contingency.

But do people actually exhibit sensitivity to blockwise prime-target contingency in priming tasks in this way? With visible primes, a recent study by Kinoshita and Norris (2010) showed that they do. Kinoshita and Norris used the same-different task, in which the response required to the target is whether it is the same (ignoring the difference in case) or different from a reference item presented just before the target (e.g., reference item = faith, target = FAITH (SAME response), or target = REPLY (DIFFERENT response)). This task shows robust masked priming effects for the SAME responses (and not for DIFFERENT response, see Norris & Kinoshita, 2008 for an explanation). Kinoshita and Norris compared the size of priming with visible primes and masked primes using two different levels of prime-target contingency. Specifically, in one condition, which was called the "predictive contingency" condition, the response to the prime was the same as that to the target on .75 of the trials and different from the target on .25 of the trials (i.e., .75 congruent and .25 incongruent); in another condition (called the "zero-contingency" condition), the response to the prime was the same as that to the target on .5 of the trials and different on .5 of the trials, i.e., the prime was not diagnostic of the response to the target. These two contingency conditions may be viewed as having high (blockwise) prime utility, and low(est) prime utility, respectively. The results showed that whereas the masked priming effect remained large and unchanged in the two contingency conditions (see also Perea & Acha, 2009, regarding a similar finding with "transposed-letter" priming effect), the priming effect with visible primes was much larger in the predictive contingency condition and reduced to a negligible level in the zero-contingency condition. In other words, when the prime was visible, but not when it was masked, priming effects varied in size as a function of blockwise prime-target contingency. Of course, the fact that with visible primes the size of priming varied with congruency proportion is unsurprising. What should be noted, however, is that in the zero-contingency condition, priming with visible primes was virtually absent, consistent with the idea that the size of priming effect with visible primes tracked the *prime utility*.

Still, it is possible that prime utility impacts differently on different tasks, and the absence of priming effect (with visible primes) in the zero-contingency condition observed by Kinoshita and Norris (2010) using the same-different task may be specific to this task. Further, our argument is that primes are diagnostic in both low- and high-proportion congruent conditions, but it is not known whether with visible primes individuals show a reverse congruence effect in a low-proportion-congruent block (Kinoshita & Norris, 2010, did not include this condition). We therefore conducted Experiments 1 and 2 to test these predictions, using visible primes in a parity decision task. Experiment 1 used a low-proportion-congruent block (containing 80% incongruent trials and 20% congruent trials), and Experiment 2 used a zero-validity block (containing 50% congruent trials and 50% incongruent trials). If subjects are sensitive to prime utility, operationalized in terms of the prime-target contingency, then a reverse congruence effect should be observed in Experiment 1, and no congruence effect should be observed in Experiment 2.

Experiment 1

Experiment 1 used visible primes in a low-proportion-congruent (i.e., high-proportion-incongruent) block. We also asked the participants after the experiment about their perception of the proportion of congruent vs. incongruent trials, in order to gauge whether the size (or direction) of the congruence effect is related to the awareness of congruency proportion.

Method

Participants. Nineteen undergraduate psychology students from Macquarie University participated in Experiments 1 for course credit.

Design. The experiment had two within-subject factors, Congruence (congruent vs. incongruent) and Block (1-3). Each block contained 120 trials (24 congruent trials and 96 incongruent trials). The dependent variables were decision latency and error rate.

Materials. The stimulus materials used in this and subsequent experiments were number words ONE to NINE, excluding FIVE, so that there were an equal number of even and odd targets. In the congruent trials, an odd target was paired with an odd prime (e.g., seven-ONE; nine-THREE), and an even target was paired with an even prime (e.g., two-FOUR, eight-SIX), but never the target itself (i.e., there were no identity/repetition primes). Thus, each target could be paired with three congruent primes (e.g., three-ONE, seven-ONE, nine-ONE). In the incongruent trials, an odd target was paired with an even prime (e.g., eight-ONE) and an even target was paired with an odd prime (e.g., one-FOUR). Each target was paired with three incongruent primes (e.g., two-ONE, four-ONE, six-ONE). This generated 24 unique congruent prime-target pairs (12 odd targets and 12 even targets) and 24 unique incongruent prime target pairs (also 12 odd targets and 12 even targets). A block consisted of 24 unique congruent trials presented once and 24 unique incongruent trials presented 4 times (120 trials).

Prior to the test proper, participants were given 20 practice and warm-up items that were representative of the block proportion. These items were not included in the analysis.

Apparatus and Procedure. Participants were tested in groups of 1-3, seated approximately 40 cm in front of a CRT monitor. Each participant completed three blocks of 120 test trials (i.e., a total of 360 trials) with a self-paced break between blocks. Within each block, a different random order of trials was generated for each participant.

Participants were instructed at the outset of the experiment that on each trial, they would be presented with two number words in succession, the first one in lowercase letters and the second in uppercase letters, and their task was to decide for the second item whether it was even or odd, as fast and accurately as possible. No mention was made of the relationship between the prime and target. They were instructed to press a button marked "+" on a response keypad for "even" with their right hand, and a key marked "-" for "odd" with their left hand.

Stimulus presentation and data collection were controlled by the DMDX display system developed by K.I. Forster and J.C. Forster at the University of Arizona (Forster & Forster, 2003). Stimulus display was synchronized to the screen refresh rate (13.3 ms).

Each trial started with the presentation of a prime (e.g., one), presented in lowercase letters in Courier 10 point font, in the center of the screen for 306 ms followed by a blank screen for 253 ms (i.e., prime-target SOA of 559 ms). This was in turn followed by the target presented in uppercase (e.g., TWO), which remained on the screen for a maximum of 2000 ms, or until the participant's response. Participants were given a feedback ("Wrong response") only following an error.

Following the experiment, participants were given a written questionnaire that asked about their subjective sense of proportion of prime-target congruency. It asked participants to choose one out of the following four alternatives, whether they thought that there were: 1) more, 2) an equal number of, or 3) fewer, congruent trials than incongruent trials, or 4) they didn't know.

Results

In Experiment 1, we carried out a conventional analysis of RT and error rates aggregated over trials, treating subject as a random factor. In Experiment 1, the preliminary treatment of trials was as follows. Any trial on which a participant made an error was excluded from the analysis of RT. To reduce the effects of extremely long and short RT, a cutoff was set for each participant at 3 S.D. units from each participant's mean RT and RTs shorter or longer than the cutoff was replaced with the cutoff value. In Experiment 1, this affected 1.7% of trials. Mean decision latencies and error rates are presented in Table 1.

Insert Table 1 about here

We report analyses of correct decision latency (RT) and error rates via a two-way analysis of variance (ANOVA) with the factors *Congruence* (congruent vs. incongruent), and *Block* (1-3). Both were within-subject factors. Effects were considered to be significant at the .05 level.

RT. The main effect of *Congruence* was significant, $F(1, 18) = 7.24$, $MSe = 4509.55$. This effect reflected a slower response for the congruent trials (574 ms) than the incongruent trials (562 ms), i.e., a *reverse* congruence effect. Congruence interacted with Block, $F(1,18) = 7.10$, $MSe = 410.96$. The interaction reflected the fact that the reverse congruence effect was largest in the last block (5 ms in Block 1, 3 ms in Block 2, 29 ms in Block 3). Separate analyses showed that the reverse congruence effect was non-significant for the average of Blocks 1 and 2, $F < 1.0$, but significant in Block 3, $F(1,18) = 24.73$, $MSe = 342.35$.

Error rate. The main effect of *Congruence* was significant, $F(1,18) = 14.39$, $MSe = 26.73$. Consistent with the latency data, this reflected worse performance for the congruent trials (8.7%) than the incongruent trials (5.0%). None of the other main or interaction effects reached significance, all $F < 1.0$.

Proportion awareness questionnaire. Nine out of 19 participants correctly answered that there were more incongruent trials than congruent trials; 2 answered there were more congruent than incongruent trials, 4 answered there were an equal number, and 4 answered they didn't know. Thus, only half of the participants were aware of the correct proportion at the end of the experiment. The nine participants who correctly identified the block as containing more incongruent trials than congruent trials were classified as "Aware", and the remaining 10 were classified as "Unaware".

Insert Figure 1 about here

Figure 1 shows the size of reverse congruence effect (a positive value indicating slower/more error prone response for the congruent trials) separately for the two groups of participants. It can be seen that both the Aware and Unaware participants showed a reverse congruence effect. A 2 (Awareness) x 2 (Congruence) x 3 (Block) ANOVA showed no significant main or interaction effect with the Awareness factor (for latency data, all $F < 2.84$, $p > .11$; for the error rate data, all $F < 3.01$, $p > .10$).

Discussion

The main result of Experiment 1 using visible primes was that in a low-proportion-congruent block (20% congruent and 80% incongruent trials), a *reverse* congruence effect was found, i.e., subjects were faster to classify a target as odd or even following a visible prime that was opposite in parity. Of course, the fact that with visible primes, people can make use of the fact that the prime is predictive of the opposite parity itself is not surprising. Neely (1977) has demonstrated a similar finding using the "shift" manipulation. Subjects were presented with three category primes ("bird", "body", "building"), and for two of the primes, they were told to expect exemplars from a different category (e.g., expect a part of body when the prime was "building"). Provided that the prime-target SOA was sufficiently long, Neely found that the response to semantically unrelated, but expected target (e.g., building – LEG) showed facilitation, and the semantically related, but unexpected target (e.g., building – DOOR) showed inhibition relative to a neutral prime "XXXX". There is an important difference between Neely's finding and the present reverse congruence effect, however.

Neely (1977) explicitly instructed the subjects to shift their expectation to pre-specified semantically unrelated categories, and told them which two of the three primes were such "shift" primes. Further, subjects were excluded if they "could not describe midway through the session the attention-shifting strategy she was to be adopting" (p.236). Subjects in Neely's (1977) experiments were therefore fully aware of the prime-target contingency. In contrast, participants in the present experiment were not informed that the prime was predictive of the opposite parity, and in a post-experiment questionnaire, only half of the participants correctly indicated the actual congruency proportion at the end of the experiment. Moreover, the size of reverse congruence effect was unrelated to the awareness of the congruency proportion, suggesting that the prime-target relationship here was learned “implicitly”.

Implicit contingency learning (learning of relationship between two events in the absence of the awareness of the relationship) has been demonstrated in other studies. For example, Schmidt, Crump, Cheesman, and Besner (2007) have shown that response in a Stroop color classification task was faster for neutral distractor words (e.g., MOVE, GRIP) when the word was presented in the color with which it was paired more frequently. This was found even in the subjects who were subjectively unaware of the relationship between the colors and words. Of note, Schmidt et al. (2007) suggested that the proportion effect observed with masked primes reported by Bodner and Masson (2001, 2003) also reflected implicit contingency learning. The present finding of *reverse* congruence effect with visible primes, in contrast to the *positive* congruence effect found with masked primes in a low-proportion-congruent block (Bodner & Dypvik, 2005) is at odds with this suggestion. In sum, the results of Experiment 1 showed that with visible primes, the priming effect tracks blockwise prime utility (operationalized as the prime-target contingency) implicitly. However, the pattern is very different from that produced by masked primes.

Experiment 2

In Experiment 2, the blocks contained an equal proportion of congruent trials and incongruent trials. We compared priming with visible primes and masked primes. With masked primes, we expected to replicate the previous findings (e.g., Reynvoet, et al., 2002, Fabre & Lemaire, 2006) of a positive congruence effect. In contrast, with visible primes, from the view that subjects are sensitive to prime utility, it was expected that there would be no congruence effect.

Method

Participants. Thirty-eight undergraduate psychology students from Macquarie University participated in Experiments 2 for course credit. Half were assigned to the visible prime condition, and the other half, to the masked prime condition, in the order of arrival.

Design. The experiment had two within-subject factors, Congruence (congruent vs. incongruent) and Block (1-3), and one between-subject factor, Prime visibility (visible vs. masked). Each block contained 96 trials. The dependent variables were decision latency and error rate.

Materials. The stimulus materials used in this experiment were the 24 congruent and 24 incongruent prime-target number word pairs used in Experiment 1. Each block contained 96 trials, composed of 24 unique congruent pairs presented twice and 24 unique incongruent pairs presented twice.

Prior to each test block, participants were given 20 practice and warm-up items that were representative of the block proportion. These items were not included in the analysis.

Apparatus and Procedure. The apparatus used and the general procedure were identical to those of Experiment 1. The trial sequence in the visible prime condition was identical to that of Experiment 1. In the masked prime condition, the trial sequence was as follows. Each trial started with the presentation of a forward mask consisting of five hash marks (#####), presented in Courier 10 point font, in the center of the screen for 506 ms. The forward mask was replaced by the prime presented in lowercase for 53 ms, which was in turn replaced by the target presented in uppercase. (The interval between the onset of forward mask and the target - 559 ms - was therefore the same as the prime-target SOA for visible primes.) The target remained on the screen for a maximum of 2000 ms, or until the participant's response. Participants were given a feedback ("Wrong response") only following an error.

Results and Discussion

The preliminary treatment of data was identical to Experiment 1. In Experiment 2, the 3 S.D. cutoff procedure affected 1.9% of trials. Mean decision latencies and error rates are presented in Table 2.

Insert Table 2 about here

We first report analyses of correct decision latency (RT) and error rates as a three-way ANOVA with the factors *Prime visibility* (Visible vs. Masked), *Congruence* (congruent vs. incongruent), and *Block* (1-3, linear and quadratic trends). Prime visibility was a between-groups factor; the other factors were within-subject factors. Where there is an interaction involving *Prime visibility*, we report separate analyses for the Visible prime group and the Masked prime group. Effects were considered to be significant at the .05 level.

RT. The linear component of the Block effect was significant, $F(1, 36) = 17.67$, $MSe = 1264.51$, showing that RT reduced linearly across blocks. The main effect of Congruence was significant, $F(1, 36) = 44.34$, $MSe = 246.65$. Critically, Congruence interacted with Prime visibility, $F(1, 36) = 48.82$, $MSe = 246.65$. As can be seen in Table 2, there was no congruence effect in the Visible prime group: $F < 1.0$, but a robust congruence effect in the Masked prime group, $F(1,18) = 101.36$, $MSe = 226.56$.

Error rate. The only significant effect was the interaction between Congruence and Mask type, $F(1,36) = 32.12$, $MSe = 13.13$. As can be seen from Table 2, this reflected the fact that whereas the Visible prime group showed a reverse congruence effect, $F(1,18) = 12.70$, $MSe = 11.52$, the Masked prime group showed a congruence advantage, $F(1,18) = 19.49$, $MSe = 14.73$.

The main result of Experiment 2, using an equal proportion of congruent and incongruent trials, was a dissociation between the effects of visible vs. masked primes. Replicating previous studies (e.g., Reynvoet, et al., 2002; Fabre & Lemaire, 2005), a robust congruence effect was found with masked primes. In contrast, with visible primes, no congruence effect was observed in RT; and there was even a reverse congruence effect with error rate. The results are consistent with our analysis that with visible primes, the priming effect tracks prime-target contingency. Prime utility is the lowest in the zero-contingency block which contains an equal proportion of congruent and incongruent trials, because the prime parity is not diagnostic of the response to the target. In this case, it is optimal to ignore the prime parity, resulting in no congruence effect. In the terms used by Huber, Shiffrin, Lyle, and Ruys (2001) in their ROUSE (Responding Optimally with Unknown Sources of Evidence) model of priming, it is optimal to “discount” the prime parity when it is undiagnostic. The reverse congruence effect observed with error rate is consistent with the possibility that subjects may have “overdiscounted” prime information (Huber, et al., 2001). In contrast, with masked primes, because participants are unaware that the prime is a separate perceptual event from the target, there is no possibility of strategically discounting the prime information and the evidence accumulated from the prime is (mistakenly) combined with the evidence for the target, hence resulting in a congruence effect.

Rational analysis of memory and masked priming. The view that the congruency proportion effect with masked primes reflects adaptation to blockwise prime utility—which was not supported by the data in Experiment 2—implies that learning is possible when the elements that constitute the relationship to be learned are not consciously identified. Bodner, Masson and Richard (2006) suggested this is possible, arguing that the relatedness proportion effect with masked primes is consistent with Anderson’s rational analysis of memory (Anderson & Milson, 1989; Anderson & Schooler, 1991). Anderson and Milson (1989) suggested that “memory is using the statistics derived from past experience to predict what memories are currently relevant” (p.703) and showed that the effects of, for example, frequency and recency, on the speed of target retrieval can be predicted on the assumption that memory is estimating which knowledge will be needed from past statistics about interitem associations (i.e., “need probability”). Bodner et al. (2006) applied this argument to masked priming, and argued that the proportion effect can be explained in terms of the need probability being higher in the high-proportion-congruent block than in the low-congruent proportion block.

We believe this application of Anderson and Milson's (1989) rational analysis framework to masked priming is mistaken. With visible primes, the statistics of prior occurrence can be kept. In contrast, with masked primes, participants are unaware that the prime and target are separate perceptual events, and hence statistics of the prime as distinct from the target cannot be kept. Humphreys, Besner and Quinlan (1988) were the first to make this distinction, and suggested that "an explicit episodic record of the prime" (p.63) capable of supporting long-term priming is formed only with visible primes. Norris and Kinoshita (2008) also made this distinction in proposing a Bayesian theory of masked priming. They argued that "A supraliminal word prime might have a genuine effect on the probability of encountering that word again in the experiment, so the prime should be expected to alter the priors of the target. .. . With masked primes, the priors are altered as a result of partial processing of the perceptual evidence for the prime. Under the circumstances of masked priming, the evidence from the prime thus leads to a revision of priors that has nothing to do with the probability of encountering the target" (p.439). What Norris and Kinoshita suggested is that with masked primes, because subjects are unaware that the prime is a separate event from the target, the evidence accumulated from the prime is combined with the evidence accumulated from the target that is used to make the decision required by the task. That is, although people approximate optimal Bayesian decision makers, masked priming reflects a limitation to this process. To borrow the term used by Huber, et al.'s (2001) model of priming, people fail to "discount" the evidence accumulated from a masked prime optimally in line with prime-target contingency (blockwise prime utility). Thus, contrary to Bodner et al. (2006), we argue that the absence of conscious identification of the prime presents a fundamental theoretical difficulty for applying the rational analysis framework of memory to masked priming. The dissociation observed here between the visible and masked primes in both the zero-contingency block and the low-proportion-congruent block is consistent with our view. Our view is also consistent with the fact that a number of studies have not been able to obtain the relatedness proportion effect with masked primes (e.g., Brysbaert, 2001; Grossi, 2006; Pecher, Zeelenberg, & Raaijmakers, 2002; Perea & Rosa, 2002).

The dissociation between the visible and masked primes observed in Experiments 1 and 2 suggests that the congruency proportion effect with masked primes does not reflect the adaptation to the blockwise prime utility. What then could be driving the proportion effect with masked primes? We next present an account that, like Bodner et al.'s, suggests a type of adaptation, but the adaptation is not to congruency proportion per se but rather to the history of trial difficulty.

The ASE account. In our previous paper (Kinoshita, et al., 2008), we explained the proportion effect observed with masked primes in a word naming task in terms of an *Adaptation to the Statistics of the Environment* (ASE) theory of optimal response initiation. The theory was proposed originally to account for sequence and blocking effects found in a wide range of speeded naming and choice tasks. The former refers to the finding that reaction time (RT) to a target is faster when the previous trial was easy than hard (e.g., Kiger & Glass, 1981; Taylor & Lupker, 2001), and the latter to the finding of RT homogenization – the fact that the difference between the easy and difficult items is smaller when they are intermixed than when they are presented in separate blocks composed purely of easy items and difficult items (e.g., Lupker, Brown & Colombo, 1997; Lupker, Kinoshita, Coltheart, & Taylor, 2003; Rastle, Kinoshita, Lupker, & Coltheart, 2003). These effects are found across many domains, e.g., in reading aloud words that do and do not follow the regular spelling-to-sound mapping rules (e.g., dot vs. yacht), naming the sum in easy vs. hard addition problems (e.g., “10 + 7 = ?” vs. “8 + 9 = ?”), making lexical decisions to high- vs. low-frequency words. It is found even when different tasks, such as naming words and pictures, or naming letter strings and making lexical decisions, are interleaved (e.g., Kiger & Glass, 1981; Lupker, et al., 2003).

The ASE explains these phenomena in terms of a mechanism of response initiation that adapts to the history of trial difficulty. Detailed description of the model’s workings is given in our published works (Kinoshita & Mozer, 2006; Kinoshita, Forster, & Mozer, 2008; Mozer, Kinoshita & Davis, 2004; Mozer, Kinoshita & Shettel, 2007; Jones, Mozer, & Kinoshita, 2009). Here we summarize the key aspects of the model that predict the proportion effect.

In speeded response tasks, subjects are instructed to respond as quickly as possible without making many errors. RT and error rate in these tasks are well-described by models of evidence accumulation, the best-known example of which is the diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008). This model is readily extended to tasks with more than two response alternatives by assuming accumulators that integrate evidence for each candidate response. In these models, trial difficulty is viewed in terms of the target drift rate—the rate at which evidence supporting the correct response accumulates over time. For easy items, evidence accumulates faster, and for hard items, evidence accumulates more slowly. (In addition to the target drift rate, the drift rates for alternative responses also matter, but we will assume they are negligible in this discussion.)

Recent theories have cast response initiation in terms of optimality (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Jones, Mozer, & Kinoshita, 2009; Mozer, Colagrosso, & Huber, 2002), where optimal responding might be defined as minimizing a linear combination of RT and error rate, or maximizing reward per unit time. According to these theories, responses initiated early risk errors, and responses initiated late incur a cost of unnecessary delay, and the optimal point of responding balances these two considerations. If the cost of responding depends on the expected accuracy of a response as a function of time, knowledge of the drift rate of the target is required. This rate varies from trial to trial, based on variability of item difficulty. One might imagine that the drift rate could be estimated as evidence builds up, but evidence accumulation is noisy and there is a chicken and egg problem: determining the drift rate of the target depends on knowing which accumulator trace corresponds to the target (assuming that there are multiple responses, and therefore multiple accumulator traces). Although this drift rate cannot be reliably determined on the fly, the priors on this drift rate can be exploited. That is, the distribution over target drift rates (e.g., the mean and variance of the distribution if it is Gaussian) can be used to estimate the cost of responding, and effectively, how the response criterion should be set.

In the specific proposal of Jones et al. (2009), the target drift-rate distribution is approximated by updating estimates of the mean and variance based on recent trials. This distribution of drift rates is used to interpret the accumulation of evidence on the current trial, and influences the evidence threshold required to initiate a response. Although the drift rate on the current trial is not affected directly by recent trials, the inferred drift rate, i.e., the estimate of the drift rate used to determine when sufficient evidence has arrived, is influenced by recent trials.

The ASE model thus predicts that performance on the current trial should be influenced by difficulty of recent trials. If recent trials have been easy, an expectation arises that the target drift rate will be high and therefore the noise from the diffusion process will be less critical. Consequently, a confident response can be made with less evidence favoring one response relative to another, and responses will tend to speed up. The reverse pattern occurs if recent trials have been difficult.

The effect of the list-wide statistics of a block, as studied in blocking paradigms, is a natural consequence of sequential effects. In a block consisting purely of easy trials, the mean drift rate is larger than the mean drift rate in a block consisting purely of hard trials, and the mean drift rate in a mixed block is necessarily intermediate between the two. Consequently, the expected drift rate of an easy item is lower in a mixed block than in a pure easy block, and the expected drift rate of a hard item is higher in a mixed block than in a pure hard block. These expected drift rates explain the RT homogenization in a mixed block (i.e., the blocking effect): Easy items will responded to at a later point in time, and the hard items will be responded to at an earlier point in time when they are mixed randomly, relative to the respective pure blocks.

In many studies using the naming task, blocking effects have been symmetric: That is, relative to pure blocks of easy and hard items, the speedup of hard items in a mixed block is approximately equal to the slowdown of easy items (see the left panel of Figure 2.) The ASE model can produce either symmetric or asymmetric blocking effects; however, when the effects are asymmetric, the ASE model predicts that easy items slow down in a mixed block more than hard items speed up, consistent with direction of asymmetry in blocking effects typically observed (see the right panel of Figure 2). In Kinoshita et al. (2008, Appendix A), we presented a mathematical proof that the condition yielding an asymmetric blocking effect also produces a congruence proportion effect. In brief, the condition requires that the easy item RTs be more sensitive to the history of trial difficulty than the hard item RTs.

The ASE model is agnostic as to whether blocking effects will be symmetric or asymmetric, but it does allow us to offer an intuition about why, when blocking effects are asymmetric, easy items are more sensitive to historical difficulty than hard items. For any item, responding later will decrease the chance of an error, because the longer evidence is allowed to accumulate, the more reliable the response should be. Likewise, responding sooner will increase the error rate. However, delaying a response by a fixed amount of time should in general benefit an easy item more than a hard item because by definition, more evidence should accumulate in the fixed time for the easy response. Likewise, responding earlier by a fixed amount of time should in general have a smaller impact on the error rate for a hard item than an easy item. Thus, the slow down of an easy item by a fixed time will reduce the cost of responding more than the speed up of a hard item will increase the cost. An optimal response initiation process therefore may manifest this asymmetry—or more slow down of the easy than speed up of the hard—as it attempts to consider the trade off between response delays and increased errors in an environment of uncertainty concerning item difficulty.

Reanalysis of Experiment 2. The foregoing section described the core claim of the ASE model, namely, that in order to respond optimally in a noisy evidence accumulation framework with uncertainty concerning item difficulty, a rational decision process will be influenced by the drift rates observed on recent trials (the difficulty history). This leads to the prediction that RT of current trial (trial N) should be sensitive to the difficulty of previous trial (trial N-1). In addition, according to the ASE model, the key condition that produces a proportion effect is that the difficulty history has a greater impact on the easy items than on the hard items. In masked priming, congruent trials are easier—they are both responded to faster and are less error prone—than the incongruent trials. Thus, a condition for observing a proportion effect according to the ASE model is that the congruent trials are more sensitive to the history of trial difficulty than the incongruent trials. To test whether this condition is met, we re-analyzed Experiment 2 using the linear mixed effects (lme) model, which allows analysis of data at the individual trial level without aggregating the data over conditions. The preliminary treatment of RT data for this analysis was as follows. First, we examined the shape of RT distribution for correct trials (a total of 5235 observation), and applied an inverse transformation (1/RT) to approximate a normal distribution, in order to meet the distributional assumption of the linear mixed model. (We used the inverse transformation rather than the log transformation because the inverse transformation approximated the normal distribution better.) We excluded trials with RTs shorter than 300 ms (5 data points). This cutoff was determined by inspecting the Q-Q plots of inverse-transformed RT. The dependent variable used in our analysis was "invRT", defined as -1000/RT: We multiplied 1/RT by -1000 to maintain the direction of effects (so that a larger invRT meant a slower response), and to avoid too many decimal places. *Lme4* (Bates, Maelchler, & Dai, 2008) and *languageR* packages (Baayen, 2008) as described in Baayen (2008) implemented in R (R Development Core Team, 2008) were used.

The statistical model we tested included two fixed factors, Congruence (congruent vs. incongruent), and prevRT (RT on previous trial),[2] and their interaction, and subjects (19) and stimuli (48 unique prime-target combination) as random factors (invRT ~ Congruence * CprevRT + (1|subject) + (1|stimuli)). We centered prevRT, or CprevRT defined as prevRT – mean(prevRT), to avoid a spurious correlation between the intercept and slope (see Baayen, 2008, p.254). To examine the effect of previous trial RT meant that the trials on which an error was made on the previous trial were excluded from analysis. There were 5007 data observations.

Insert Table 3 about here

Table 3 summarizes the coefficients of the fixed effects of the model (for the Visible prime group and the Masked prime group separately), together with their standard error, and the t-values.[3] The model shows that the effect of Congruence is significant in

[2] The analysis included only the correct trials.

[3] As noted by Baayen (2008, p.247), the current version of lme4 package does not provide p-values for t-tests, as it is presently unclear how to calculate the appropriate degrees of freedom. However, as also noted by Baayen, for large number of observations (> 100), it is safe to regard |t| > 2.0 as significant at the .05

the Masked prime group ($t = 4.80$, $p < .0001$) but not in the Visible prime group ($t = -.31$, $p > .75$), consistent with the conventional analysis of RT aggregated over trials. It can also be seen that CprevRT has a large effect for both groups ($t = 5.77$, $p < .0001$ in the visible prime group; $t = 11.5$, $p < .0001$ in the masked prime group). The positive value of the coefficient means that as CprevRT increased, invRT (-1000/RT) increased, that is, current trial RT is positively correlated with previous trial RT. The critical finding is that the effect of CprevRT interacted with Congruence in the Masked prime group ($t = -2.00$, $p < .05$), though not in the Visible prime group ($t = -.30$, $p > .76$). Figure 3 visualizes the effect of CprevRT on the current trial RT, with the masked prime group on the left panel and the visible prime group on the right panel. Note that in generating the plot, the dependent variable invRT (-1000/RT) was transformed back to RT (and hence the functions relating the RTs on the current trial and the previous trial are curved, rather than linear). The curves correspond to the function relating previous trial RT to current trial RT, estimated by the equation invRT ~ Congruence * CprevRT + (1|subject) + (1|stimuli). It can be seen that the slope for the congruent trials is greater than the incongruent trials in the masked prime group, indicating that the precondition for finding a proportion effect according to the ASE model is met.

Insert Figure 3 about here

----

level. P-values reported here were estimated using the Markov Chain Monte Carlo (MCMC) sampling method as implemented in the l*anguageR* package (Baayen, 2008).

Experiment 3

The foregoing analysis of Experiment 2 provided support for the core claim of ASE model in showing that previous trial difficulty, indexed here by previous trial RT, affects the current trial RT. It also showed that the congruent trials (the easy trials) are more sensitive to previous trial RT than the incongruent trials, thus satisfying the condition for expecting a proportion effect according to the ASE model.

Having established that the key condition for expecting a proportion effect was met, Experiment 3 tested whether the ASE model can provide a viable account of the proportion effect in masked priming using Bodner and Dypvik's (2005) proportion manipulation. In the high-proportion-congruent block, 80% of trials were congruent and the remaining 20% were incongruent; in the low-proportion-congruent block, 20% of trials were congruent and the remaining 80% were incongruent. We expected to replicate Bodner and Dypvik's (2005) finding of proportion effect, with a larger congruence effect in the high-proportion-congruent block than the low-proportion-congruent block. The critical prediction concerned interactions involving previous trials. First, according to the ASE model, for a proportion effect to be observed, the easy (congruent) trials need to be more sensitive to previous trial difficulty than the hard (incongruent) trials, i.e., the interaction between congruence and previous trial RT observed in Experiment 2 is expected. Second, the ASE model claims that what is critical about the congruency proportion manipulation is not the block type per se, i.e., high versus low proportion, but rather the fact that in a high congruency block, many of the previous trials are easy, and in a low congruency block, many are hard. Consequently, even analyses conditioned on trial N-1 may show an effect of blocks because the block determines the statistics of trials

N-2, N-3, etc. In general, a high proportion block will produce a prior distribution over item difficulty that is biased toward high drift rates, and a low proportion block will produce a distribution biased toward low drift rates. If the asymmetry we have previously discussed is present—the asymmetry that leads to more sensitivity of easy items to trial history than hard items—the asymmetry might produce an interaction due to block composition as well: In a high proportion block, where the inferred drift rates will be higher and therefore the items will be treated as easier, sequential effects should be stronger. Consequently, the nature of the interaction of the previous and current trials may depend on the block composition. In sum, the ASE account predicts that for a proportion effect to be observed, the effect of previous trial RT should interact with the effect of difficulty of current trial (congruence), and there may also be a triple interaction between effects of previous trial RT, the difficulty of current trial (congruence), and the proportion of easy trials (congruency proportion).

Method

Participants. Thirty-eight Macquarie University students participated in Experiment 3 in return for course credit.

Design. The experiment had two within-subject factors, Congruence (congruent vs. incongruent) and Block (1-3), and one between-group factor, Proportion (high-proportion-congruent vs. low-proportion-congruent). The dependent variables were decision latency and error rate.

Materials. The stimulus materials used in this experiment were the 24 congruent and 24 incongruent prime-target number word pairs used in the previous experiments.

Each block contained 3 subblocks of 120 trials each. In each subblock, in the high-proportion-congruent condition, 96 trials were congruent and 24 trials were incongruent; in the low-proportion-congruent condition, 24 trials were congruent and 96 trials were incongruent.

Prior to each test block, participants were given 20 practice and warm-up items that were representative of the block proportion. These items were not included in the analysis.

Apparatus and Procedure. The apparatus used and the general procedure were identical to those of the Masked prime condition in Experiment 2.

Results

We report two sets of analyses, one using the conventional ANOVA with data aggregated over trials in a condition, and one using the linear mixed effect model. The former analysis is to maintain comparability with previous studies, and to establish that the proportion effect is present. The latter allows us to examine the effect of previous trial difficulty at the level of individual trials.

Analysis based on aggregated data. The preliminary treatment of data was identical to previous experiments, and the 3 S.D. cutoff procedure affected 1.9% of trials. We analyzed the mean RT from correct trials and mean error rate using a three-way ANOVA with the factors *Congruence* (congruent vs. incongruent), *Proportion* (high- vs. low proportion of congruent trials), and *Block* (1-3). The proportion effect is indicated by the *Congruence* x *Proportion* interaction. *Proportion* was a between-group factor; the other two factors were within-subject factors. Effects were considered to be significant at the .05 level. The summary data are presented in Table 4.

Insert Table 4 about here

RT. The main effect of *Congruence* was significant, $F(1,36) = 52.71$, $MSe = 686.85$: Congruent trials were 25 ms faster than incongruent trials. Importantly, *Congruence* and *Proportion* interacted, $F(1,36) = 9.77$, $MSe = 686.85$, demonstrating a proportion effect. Congruence effect was greater in the high-proportion group (36 ms) than the low-proportion group (14 ms). There was a significant linear trend of Block, $F(1,36) = 18.86$, $MSe = 1261.74$. None of other main or interaction effects reached significance, $F < 2.45$, $p > .13$ in all cases.

Error rate. The main effect of *Congruence* was significant, $F(1,36) = 3.16$, $MSe = 18.45$: Congruent trials were 3.5% *more* accurate than incongruent trials. *Congruence* and *Proportion* interacted, $F(1,36) = 4.51$, $MSe = 18.45$, demonstrating a proportion effect. Congruence effect was greater in the high-proportion group (4.7%) than the low-proportion group (2.3%). There was a significant linear trend of *Block*, $F(1,36) = 4.48$, $MSe = 23.33$. There were no other main or interaction effect that reached significance, all $F < 1.60$, $p > .21$.

Linear mixed effect model of RT. The analysis we report is based on RTs from correct trials, excluding RTs shorter than 300 ms. The preliminary treatment of RT was identical to the reanalysis of Experiment 2, and we used -1000/RT (invRT) as the dependent measure. As in Experiment 2, when an error was made on the previous trial, it was excluded from analysis. There were 12380 data observations.[4]

[4] Due to experimenter error, in the low-proportion-congruent condition, one of the 24 congruent primes used an identity prime (nine – NINE). 52 data observations were lost when these trials were excluded.

The statistical models we tested included three fixed factors, Congruence (congruent vs. incongruent), Proportion (hi- vs. low proportion of congruent trials) and CprevRT (RT on previous trial, centered), and subjects (38) and stimuli (48 unique prime-target combination) as random factors. We tested two models, a full model involving a triple interaction between the three fixed factors, and another without the triple interaction involving CprevRT. Comparison of the two models showed that the more complex model involving the triple interaction provided a better fit to the data, χ^2 (2) = 24.80, $p < .001$, hence this is the model we report here (invRT ~ Congruence * Proportion * CprevRT + (1|subject) + (1|stimuli)).

Insert Table 5 about here

Table 5 summarizes the coefficients of the fixed effects of the model, as well as the standard error and t-values. It shows that the effect of Congruence is significant ($t = 6.78$, $p < .0001$), as well as the interaction between Congruence and Proportion, i.e., a proportion effect ($t = -4.46$, $p < .0001$), consistent with the conventional analysis of RT aggregated over trials. The main effect of CprevRT is highly significant ($t = 15.84$, $p < .0001$). Critically, the two predicted interactions involving CprevRT were significant: CprevRT interacted with Congruence ($t = -3.87$, $p < .0001$), and there was also a triple interaction between CprevRT, Congruence, and Proportion ($t = 2.35$, $p < .02$). CprevRT also interacted with Proportion ($t = -2.39$, $p < .02$).

Insert Figures 4 and 5 about here

The critical predictions are shown in Figure 4 and Figure 5. Figure 4 visualizes the interaction between CprevRT and Congruence: It can be seen that the congruent trials ("C") are more sensitive to CprevRT than the incongruent trials ("IC"). Figure 5 shows the triple interaction between CprevRT, Congruence and Proportion, indicating that the congruent trials were more sensitive to the previous trial RT in the easy environment (the high-proportion congruent block = dashed lines) than in the hard environment (the low-proportion-congruent block = solid lines), but for the incongruent trials the interaction between previous trial RT and Proportion was not significant.

Discussion

Experiment 3 replicated the proportion effect with masked primes reported by Bodner and Dypvik (2005): The congruency advantage was greater in the high-proportion congruent block than in the low-proportion-congruent block. The novel finding revealed by the linear mixed effects analysis concerned the triple interaction between previous trial RT, congruence and proportion of congruent trials. As expected from the view that easy items are more sensitive to the task environment, the easy items (congruent trials) showed a greater sensitivity to the history of trial difficulty in the easy environment (high-proportion-congruent block) than in the hard environment (low-proportion-congruent block), but little difference was observed with the hard items (incongruent trials). The analysis also replicated the finding of Experiment 2 that congruent trials are more sensitive to the previous trial RT than the incongruent trials.

These two interactions are consistent with the prediction of the ASE model, which explains the proportion effect in terms of the greater sensitivity of easy items than the hard items to the difficulty of the task environment.

General Discussion

Bodner and Dypvik (2005) reported that the congruence effect in parity decision task with masked primes was reduced in a block containing a high proportion of incongruent trials relative to a block containing a high proportion of congruent trials. This effect of relatedness proportion has been taken to reflect the sensitivity of masked priming to the utility of prime information across trials. We argued that the fact that the prime is not consciously identified presents a fundamental theoretical difficulty for this view.

The present study makes two primary contributions to our understanding of the proportion effect. First, we pointed out that congruency proportion and prime-target contingency are not the same thing, and showed that visible primes (but not masked primes) showed adaptation to the prime-target contingency. Specifically, we pointed out that the low-proportion-congruent block containing 80% incongruent and 20% congruent trials has high prime utility, because the prime parity is predictive of the opposite response to the target (if the prime is even, the target is likely to be odd and vice versa). Prime utility is in fact the lowest in the zero-validity block containing an equal proportion of congruent and incongruent trials. In line with this analysis, visible primes (but not masked primes) showed a *reverse* congruence effect (advantage for incongruent trials) in the low-proportion-congruent block (Experiment 1) and *no* congruence effect in the zero-contingency block (Experiment 2). Of interest, the finding of the reverse congruence effect in Experiment 1 was unrelated to the subjective awareness of congruency proportion, indicating that the sensitivity to the prime validity here was implicit (i.e., it did not reflect a "conscious predictive strategy", cf. Neely, 1977). Nevertheless, the pattern of proportion effects observed with visible primes was qualitatively different from those observed with masked primes, indicating that the latter does not reflect the adaptation to prime utility.

The second contribution of the present study is to provide an alternative explanation of the proportion effect found with masked primes in terms of our ASE model. A core claim of this account is that because the evidence accumulation process is noisy, optimal response initiation processes need to consider the difficulty history of recent trials to determine the threshold of evidence for responding on the current trial. This predicts an effect of history of trial difficulty (indexed by the previous trial RT) on the current trial RT. Analysis of individual trial RT using the linear mixed effect model revealed a robust effect of previous trial RT, consistent with the core claim of the model. Proportion effects are predicted by the ASE when the easy items (congruent trials) are more sensitive than the hard items (incongruent trials) to the difficulty of the task environment. The linear mixed effects model analysis supported this explanation: It showed that in Experiment 3, showing a proportion effect, the congruent trials were more sensitive to the previous trial RT than the incongruent trials, and this was magnified in the easy task environment (the high-proportion-congruent block) relative to the hard task environment (the high-proportion-incongruent block). These results suggest that the ASE model offers a viable explanation of the proportion effect observed with masked primes.

When is the proportion effect absent? As noted earlier, Bodner and colleagues have reported finding the proportion effect with masked primes in several tasks,[5] including lexical decision (Bodner & Masson, 2001, 2003), naming (Bodner & Masson, 2004) and both number magnitude and parity judgment tasks (Bodner & Dypvik, 2005). Also as noted, the finding of proportion effect in the lexical decision task has been mixed. We have previously applied ASE to the naming task (Kinoshita, et al., 2008), and here to the parity judgment task. Can the ASE be applied to the lexical decision task, and can it explain why the proportion effect has sometimes been absent?

The mechanism of ASE is general across tasks, and hence there is no reason why it cannot be applied to the proportion effect in the lexical decision task. In addition, the ASE offers a coherent explanation for why the finding of proportion effect in this task (and other tasks) has been mixed.

One reason why proportion effects have been more elusive in the lexical decision task is that nonword targets are generally insensitive to masked priming. (Forster, 1998, reported that in a survey of 40 experiments, only in three cases was priming for nonword targets statistically significant at the .05 level.) This means that manipulation of difficulty of the environment as implemented by the proportion of identity-primed vs. control trials is carried by only half of the trials (word target trials). That is, the relatedness proportion manipulation (with masked primes) in the lexical decision task is *a priori* weaker than in other tasks.

[5] Also as noted earlier, a number of studies (e.g., Brysbaert, 2001; Pecher, et al., 2002) reported not being able to obtain relatedness proportion effects with masked primes, using the perceptual identification task. We have pointed out in our previous paper that this is to be expected from the ASE account (but not from the view that relatedness proportion effects reflect the prime utility) as the perceptual identification task does not require a speeded response.

Another factor is trial-by-trial variability. In the first of the series of papers on the proportion effect, Bodner and Masson (2001) noted that proportion effect was absent in some of their experiments (see their Table 1, p.633), and suggested the level of variability from trial to trial was a factor in finding a proportion effect. Specifically, proportion effect was absent when: 1) discontinuous bands of word frequency were used, 2) easy words but difficult nonwords were used; 3) difficult words but easy nonwords were used. Bodner and Masson noted that "large variation in processing difficulty (drift rate in a diffusion model) from trial to trial implies a high level of noise in the information accumulation process", and suggested that "an adaptive response to a high level of noise may be to rely heavily on recruitment of prime resources at all times, regardless of their validity" (p.633). We offer a different interpretation of the consequence of a high level of trial-by-trial variability. When trial difficulty varies wildly from trial to trial, there would be a high overlap in the distribution of drift rates between the easy items (identity-primed trials) and hard items (control trials). Consequently, the difficulty of the high-proportion-related block and low-proportion-related block would also overlap. Effectively, the difference in the task environment difficulty due to the proportion manipulation would be smaller. Just as a small numerical difference between easy and hard items would produce a weaker proportion effect, the proportion effect would be weaker the more variability exists *within* the high-proportion-related and low-proportion-related blocks.

The proportion effect has also been absent in other cases. For example, Bodner and Stalinski (2008) reported that they did not observe a proportion effect using the lexical decision task when manipulating block proportion within subjects. In the study, each subject was presented with 160 trials (80 words and 80 nonwords) of high-proportion-identity-primed block and 160 trials of low-proportion-identity-primed block. Using the parity decision task, in a separate study from those reported here, we too failed to find a proportion effect when proportion was manipulated within subjects, and each subject completed 120 trials of high-proportion-congruent block and 120 trials of low-proportion-congruent block. (In our study and in Bodner & Stalinski's study, the order of blocks was counterbalanced between subjects.) In contrast, in the present Experiment 3 and in the earlier studies by Bodner and colleagues that found proportion effects, each proportion block contained over 300 trials (360 trials in the studies using the parity decision task, and 400 trials in the studies using the lexical decision task). The failures to find statistically significant proportion effects with fewer trials suggest the possibility that an accurate estimate of the difficulty of the environment requires many trials.

To test this, we analyzed just the first block (containing 120 trials) of Experiment 3, using the same factors as the previous analysis. There was a congruence effect of 27 ms in the high-proportion-congruent block and 18 ms in the low-proportion-congruent block (refer to data in Table 4). The conventional ANOVA based on data aggregated over trials showed that the interaction between congruence and proportion (i.e., the proportion effect) was non-significant, $F(1,36) < 1.0$, indicating that 120 trials was not sufficient for a statistically significant proportion effect to be detected [6].

---

[6] Bodner and Masson (2001) analyzed data aggregated across experiments and reported that the proportion effect did not interact with block (each block containing 100 trials, see their Figure 13, p. 639). Based on this finding, they concluded "in the interim" that "sensitivity to prime validity emerges either during the 40 trials of the practice block or within the first 100 trials of the critical block"(p.638). However, they did not analyze whether the proportion effect was in fact present in the first block of 100 trials. Given that the Block x Proportion x Congruence interaction was also non-significant in our Experiment 3, their interim conclusion seems to have been premature.

The linear mixed effects model analysis involving the previous trial difficulty factor (CprevRT) informs us why a large number of trials are needed. The coefficients, standard error, and the t-values of the fixed factors of the lme model (invRT ~ Congruence * Proportion * CprevRT + (1|subject) + (1|stimuli) are shown in Table 6.

Insert Table 6 about here

It can be seen that the interaction between Congruence and Proportion is non-significant in this analysis ($t = -1.33$, $p > .18$), indicating an absence of a statistically significant proportion effect, consistent with the conventional analysis of aggregated data. Critically, unlike the analysis for the whole data set of Experiment 3, CprevRT x Congruence interaction ($t = -1.40$, $p > .16$), and the triple interaction between CprevRT, Congruence, and Proportion ($t = 0.66$, $p > .51$) were both non-significant. The effect of CprevRT was nevertheless highly robust ($t = 7.67$, $p < .0001$), indicating that current trial RT is sensitive to *immediate* trial history. What this suggests is that fine-tuning the estimated point for optimal responding as a function of global (blockwise) task environment requires many trials (> 120 trials). This is consistent with the idea that evidence accumulation is an inherently noisy process, which is a key assumption of the ASE. A practical implication of this analysis is that in order to produce a statistically significant proportion effect by manipulating the difficulty of blocks where the effect of difficulty on RT is not large, many trials (> 120 trials) are required.

Conclusion. The results presented here challenge the idea that the proportion effect observed with masked primes reflects the adaptation to prime utility across trials. Our results showed that with visible primes, people indeed adapt to the blockwise prime utility (prime-target contingency), but that pattern is different from the congruency proportion effect observed with masked primes. The latter is instead explained in terms of a mechanism of optimal response initiation, which attempts to estimate the optimal point for responding on the current trial based on the history of trial difficulty. This was supported by an analysis of the effect of RT of previous trial, using the linear mixed effects model.

In closing, we note that as the use of the linear mixed effect model, and the analysis of data at the individual trial level becomes more widespread, researchers are increasingly recognizing the impact of the previous trial RT in speeded response tasks (e.g., Kuperman, Schreuder, Bertram, & Baayen, 2009; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). Indeed, in our experience, we have not yet found a RT experiment which did not show an effect of this variable. Although our theoretical framework suggests that previous trial difficulty is the critical determinant of sequential effects, difficulty cannot be measured directly, and RT can be used as a proxy, even if the RT of the previous trial is affected in turn by the trials that came before it.

Ultimately, an understanding of trial history on response initiation processes will require more than empirical analyses of RT. It will be necessary to consider specific mechanisms and models of response initiation, and under what circumstances these theories predict that response initiation processes are modulated by recent experience. The framework we have developed, the ASE model, suggests that this modulation is a consequence of optimal decision making and should be widespread across tasks, stimuli, and responses. Other theories also characterize sequential effects as being intrinsic to decision processes (e.g.,Botvinick, Braver, Barch, Carter, & Cohen, 2001; Jones, Cho, Nystrom, Cohen, & Braver, 2002; Vickers & Lee, 1998; Wilder, Jones, & Mozer, 2009). To the extent that sequential effects are a fundamental consequence of decision processes, they will—as Kiger and Glass commented in 1981—“continue to be rediscovered in many circumstances” (p.697), and research may profit by discarding the focus on effects that depend on comparison of trial blocks with different statistics to effects that depend on how each trial or experience influences the next.

References

Anderson, N. (1960). Effect of first-order conditional probability in a two-choice learning situation. *Journal of Experimental Psychology, 59*, 73-93.

Anderson, J.R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review, 96*, 703-719.

Anderson, J.R., & Schooler, L.J. (1991). Reflections of the environment in memory. *Psychological Science, 2*, 396-408.

Baayen, R.H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Bates, D., M., Maechler, M., & Dai, B. (2008). Lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-24.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General, 129*, 177-192.

Bodner, G. E., & Dypvik, A. T. (2005). Masked priming of number judgments depends on prime validity and task. *Memory & Cognition*, *33*, 29-47.

Bodner, G. E., & Masson, M.E.J. (2001). Prime validity affects masked repetition priming: Evidence for an episodic resource account of priming. *Journal of Memory and Language, 45*, 616-647.

Bodner, G.E., & Masson, M.E.J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review, 10*, 645-652.

Bodner, G. E., & Masson, M. E. J. (2004). Beyond binary judgments: Prime validity modulates masked repetition priming in the naming task. *Memory & Cognition, 32*, 1-11.

Bodner, G.E., Masson, M.E.J., & Richard, N.T. (2006). Repetition proportion biases masked priming of lexical decisions. *Memory & Cognition, 34*, 1298-1311.

Bodner, G.E., & Stalinski, S.M. (2008). Masked repetition priming and proportion effects under cognitive load. *Canadian Journal of Experimental Psychology, 62*, 127-131.

Bogasz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J.D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*, 700-765.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624-652.

Brysbaert, M. (2001). Prelexical phonological coding of visual words in Dutch: Automatic after all. *Memory & Cognition, 29*, 765-773.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*, 284-301.

Dixon, P. & Glover, S. (2009). Perseveration and contrast effects in grasping. *Neuropsychologia, 47*, 1578-1584.

Estes, W. (1950). Toward a statistical theory of learning. *Psychological Review, 57*, 94-107.

Fabre, L., & Lemaire, P. (2005). Age-related differences in automatic stimulus-response associations: Insights from young and older adults' parity judgments. *Psychonomic Bulletin & Review, 12*, 1100-1105.

Forster, K.I. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research, 27*, 203-233.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10*, 680-698.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation and responses. *Journal of Experimental Psychology: General, 121*, 480-506.

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology, 43*, 21-30.

Huber, D.E., Shiffrin, R.M., Lyle, K.B., & Ruys, K.I. (2001). Perception and preference in short-term word priming. *Psychological Review, 108*, 149-182.

Humphreys, G.W., Besner, D., & Quinlan, P.T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General, 117*, 51-67.

Jaskowski, P., Skalska, B., & Verleger, R. (2003). How the self controls its "Automatic Pilot" when processing subliminal information. *Journal of Cognitive Neuroscience, 15*, 911–920.

Jestead, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 92-104.

Johnston, W.A., Hawley, K.J., & Elliott, J.M.G. (1991). Contribution of perceptual fluency to recognition memory judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 210-233.

Jones, A. D., Cho, R. Y., Nystrom, L. E., Cohen, J. D., & Braver, T. S. (2002). A computational model of anterior cingulate function in speeded response tasks: Effects of frequency, sequence, and conflict. *Cognitive, Affective, & Behavioral Neuroscience, 2*, 300-317.

Jones, M., Mozer, M.C., & Kinoshita, S. (2009). Optimal response initiation: Why recent experience matters. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in Neural Information Processing Systems 21*, Curran Associates, Inc. 785-792.

Jones, M., & Sieck, W. R. (2003). Learning myopia: An adaptive recency effect in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 626-640.

Kiger, J.I., & Glass, A.L. (1981). Context effects in sentence verification. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 688-700.

Kinoshita, S., Forster, K.I., & Mozer, M.C. (2008). Unconscious cognition isn't that smart: Modulation of masked repetition priming effect in the word naming task. *Cognition*, *107*, 623-649.

Kinoshita, S., & Mozer, M.C. (2006). How lexical decision is affected by recent experience: Symmetric vs. asymmetric frequency-blocking effects. *Memory & Cognition*, *34*, 726-742.

Kinoshita, S., & Norris, D. (2010). Masked priming effect reflects evidence accumulated by the prime. *Quarterly Journal of Experimental Psychology, 63*, 194-204.

Klapp, S. (2007). Nonconscious control mimics a purposeful strategy: Strength of Stroop-like interference is automatically modulated by proportion of compatible trials. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1366-1377.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society, B, 362*, 857-875. doi: 10.1098/rstb.2007.2093

Kunst-Wilson, W.R., & Zajonc, R.B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science, 207*, 557-558.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 876-895.

Kristjansson, A. (2006). Rapid learning in attention shifts: A review. *Visual Cognition, 13*, 324-362.

Laming, D. (1979). Autocorrelation of choice-reaction times. *Acta Psychologica, 43*, 381-412.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 570-590.

Lupker, S. J., Kinoshita, S., Coltheart, M., & Taylor, T. (2003). Mixing costs and mixing benefits in naming words, pictures, and sums. *Journal of Memory and Language, 49*, 556-575.

Maloney, L.T., DalMartello, M.F., Sahm, C., & Spillmann, L. (2005). Past trials influence perception of ambiguous motion quartets through pattern completion. *Proceedings of the National Academy of Sciences, 102*, 3164-3169.

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language, 48*, 131–147.

Masson, M.E.J., & Bodner, G.E. (2003). A retrospective view of masked priming: Toward a unified account of masked and long-term repetition priming. In S. Kinoshita & S.J. Lupker (Eds.), *Masked priming: The state of the art*. New York: Psychology Press, pp.57-94.

Mozer, M. C., Kinoshita, S., & Davis, C. (2004). Control of response initiation: Mechanisms of adaptation to recent experience. *Proceedings of the Twenty Sixth Annual Conference of the Cognitive Science Society* (pp. 981-986). Hillsdale, NJ: Erlbaum Associates.

Mozer, M.C., Kinoshita, S., & Shettel, M. (2007). Sequential dependencies in human behavior offer insights into cognitive control. In W. Gray (Ed.), *Integrated Models of Cognitive Systems*. Oxford University Press, pp.180-193.

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental psychology: Human Perception and Performance, 106*, 226-254.

Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp.264-336). Hillsdale: Erlbaum.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychoogy: General, 137*, 434-455.

Pecher, D., Zeelenberg, R., & Raajmakers, J.G.W. (2002). Associative priming in a masked perceptual identification task: Evidence of automatic processes. *Quarterly Journal of Experimental Psychology, 55A*, 1157-1173.

Perea, M., & Acha, J. (2009). Does letter position coding depend on consonant/vowel status? Evidence with the masked priming technique. *Acta Psychologica, 130*, 127-237.

Perea, M., & Rosa, E. (2002). Does the proportion of associatively related pairs modulate the associative priming effect at very brief stimulus-onset asymchronies? *Psicologia, 19*, 311-319 (in Spanish).

R Development Core Team (2008). R: A language and environment for statistical computing (Computer software manual]. Vienna, Austria. http://www.R-project.org

Rabbitt, P. M. A. (1989). Sequential reactions. In D.H. Holding (Ed.), *Human Skills* (2nd edition) (pp. 147-170). Chichester, U.K.: Wiley.

Rastle, K., Kinoshita, S., Lupker, S. J., & Coltheart, M. (2003). Cross-task strategic effects. *Memory & Cognition, 31*, 867-876.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59-108.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision Model: Theory and data for two-choice decision tasks. *Neural Computation, 20*, 873-922.

Remington, R. (1960). Analysis of sequential effects in choice reaction times. *Journal of Experimental Psychology, 82*, 250-257.

Reynvoet, B., Caessens, B., & Brysbaert, M. (2002). Automatic stimulus-response associations may be semantically mediated. *Psychonomic Bulletin & Review, 9*, 107-112.

Schmidt, J.R., Crump, M.J.C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition, 16*, 421-435.

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 117-138.

Vickers, D., & Lee, M. D. (1998). Dynamic models of simple judgments: I. Properties of a self-regulating accumulator module. *Nonlinear Dynamics, Psychology, & Life Sciences, 2*, 169-194.

Wilder, M., Jones, M., & Mozer, M. C. (2010).Sequential effects reflect parallel learning of multiple environmental regularities. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems, Volume 22* (pp. 2053-2061). La Jolla, CA: NIPS Foundation.

Whittlesea, B.W.A., & Price, J.R. (2001). Implicit/explicit memory versus analytic/nonanalytic processing: Rethinking the mere exposure effect. *Memory & Cognition, 29*, 234-246.

Yu, A. J., & Cohen, J. D. (2009). Sequential effects: Superstition or rational behavior? *Advances in Neural Information Processing Systems, Volume 21* (pp. 1873-1880). La Jolla, CA: NIPS Foundation.

Table 1.

Mean Response Latencies (RT, in ms) and Percent Errors (%E) in Experiment 1 (visible prime, low-proportion-congruent)

	Prime type					
	Congruent		Incongruent		Congruence effect	
Block	RT	%E	RT	%E	RT	%E
Overall	574 (13)	8.7	562 (15)	5.0	-12	-3.7
Block 1	565 (13)	9.0	560 (16)	5.4	-5	-3.6
Block 2	571 (16)	8.3	568 (18)	5.6	-3	-2.7
Block 3	587 (15)	8.8	558 (15)	4.1	-29	-4.7

Table 2.

Mean Response Latencies (RT, in ms) and Percent Errors (%E) in Experiment 2 (visible and masked primes, equal proportion congruent-incongruent)

	Prime type					
	Congruent		Incongruent		Congruence effect	
Block	RT	%E	RT	%E	RT	%E
Visible prime group						
Overall	524 (14)	6.8	523 (15)	4.6	-1	-2.2
Block 1	534 (14)	7.5	533 (14)	4.1	-1	-3.4
Block 2	523 (15)	6.5	523 (15)	5.2	0	-1.3
Block 3	515 (14)	6.5	514 (14)	4.4	-1	-2.1
Masked prime group						
Overall	530 (14)	2.8	558 (15)	5.9	28	3.1
Block 1	550 (14)	3.7	566 (14)	5.5	16	1.8
Block 2	529 (15)	2.9	562 (15)	6.9	33	4.0
Block 3	511 (14)	1.7	547 (14)	5.4	36	3.7

Table 3.

Coefficients of Congruence (C vs. IC), Previous trial RT (CprevRT, centered), and the interactions in the regression model for the inverse RTs of Experiment 2, together with the standard error (Std. Error) and the t-values

Unmasked	Estimate	Std. Error	t value
(Intercept)	-2.02E+00	4.56E-02	-44.28
Congruence	-8.08E-03	2.63E-02	-0.31
CprevRT	3.07E-04	5.31E-05	5.77
Congruence:CprevRT	-2.17E-05	7.23E-05	-0.30
Masked			
(Intercept)	-1.98E+00	3.76E-02	-52.62
Congruence	1.19E-01	2.48E-02	4.80
CprevRT	5.48E-04	4.77E-05	11.50
Congruence:CprevRT	-1.27E-04	6.33E-05	-2.00

Table 4.

Mean Response Latencies (RT, in ms) and Percent Errors (%E) in Experiment 3 (masked primes)

	Prime type					
	Congruent		Incongruent		Congruence Effect	
Block	RT	%E	RT	%E	RT	%E
High-proportion-congruent (.8 congruent, .2 incongruent)						
Overall	527 (14)	3.4	563 (12)	8.0	36	4.6
Block 1	542 916)	2.6	569 (13)	6.4	27	3.7
Block 2	521 (16)	3.2	559 (12)	8.8	38	5.6
Block 3	517 (15)	4.2	561 (13)	9.0	44	4.8
Low-proportion-congruent (.2 congruent, .8 incongruent)						
Overall	549 (15)	3.2	564 (12)	5.5	14	2.3
Block 1	565 (16)	3.3	583 (13)	5.3	18	2.0
Block 2	552 (16)	2.2	559 (13)	4.4	7	2.2
Block 3	531 (15)	4.2	549 (13)	6.8	18	2.6

Table 5.

Coefficients of Congruence (C vs. IC), Proportion (high vs. low), Previous trial RT (CprevRT, centered), and the interactions in the regression model for the inverse RTs of Experiment 3, together with the standard error (Std. Error) and the t-values

	Estimate	Std. Error	t value
(Intercept)	-1.99E+00	3.62E-02	-54.91
Congruence	1.36E-01	2.00E-02	6.78
Proportion	7.74E-02	4.91E-02	1.58
CprevRT	5.66E-04	3.57E-05	15.84
Congruence:Proportion	-6.78E-02	1.52E-02	-4.46
Congruence:CprevRT	-2.63E-04	6.78E-05	-3.87
Proportion:CprevRT	-1.79E-04	7.47E-05	-2.39
Congruence:Proportion:CprevRT	2.35E-04	9.98E-05	2.35

Table 6.

Coefficients of Congruence (C vs. IC), Proportion (high vs. low), Previous trial RT (CprevRT, centered), and the interactions in the regression model for the inverse RTs in the first block of Experiment 3, together with the standard error (Std. Error) and the t-values

	Estimate	Std. Error	t value
(Intercept)	-1.93E+00	3.88E-02	-49.83
Congruence	1.02E-01	2.59E-02	3.93
Proportion	7.24E-02	5.34E-02	1.36
CprevRT	4.74E-04	6.18E-05	7.67
Congruence:CprevRT	-1.62E-04	1.15E-04	-1.40
Congruence:Proportion	-3.33E-02	2.50E-02	-1.33
CprevRT:Proportion	-1.38E-04	1.21E-04	-1.14
Congruence:CprevRT:Proportion	1.07E-04	1.62E-04	0.66

Author notes

Sachiko Kinoshita, Macquarie Centre for Cognitive Science (MACCS) and Department of Psychology, Macquarie University, Sydney, Australia 2109. Michael C. Mozer, Department of Computer Science and Institute of Cognitive Science, University of Colorado, Boulder, CO 80309. Kenneth I Forster, Department of Psychology, University of Arizona, Tucson, AZ 85721.

Thanks are due to Shaun Greenfield and Amanda Selwood for research assistance. This study was supported by the Australian Research Council Discovery Project Grant (DP0556805) and USA National Science Foundation (project no. 0518319). We are also grateful to Mike Masson, Marc Brysbaert and an anonymous reviewer for their comments on an earlier version of the paper, and the action editor Stephen Monsell for guidance.

Mail correspondence concerning this article may be addressed to Sachiko Kinoshita, Macquarie Centre for Cognitive Science (MACCS) and Department of Psychology, Macquarie University, Sydney, NSW, Australia, 2109. Electronic mail may be sent to: sachiko.kinoshita@mq.edu.au.

List of figures

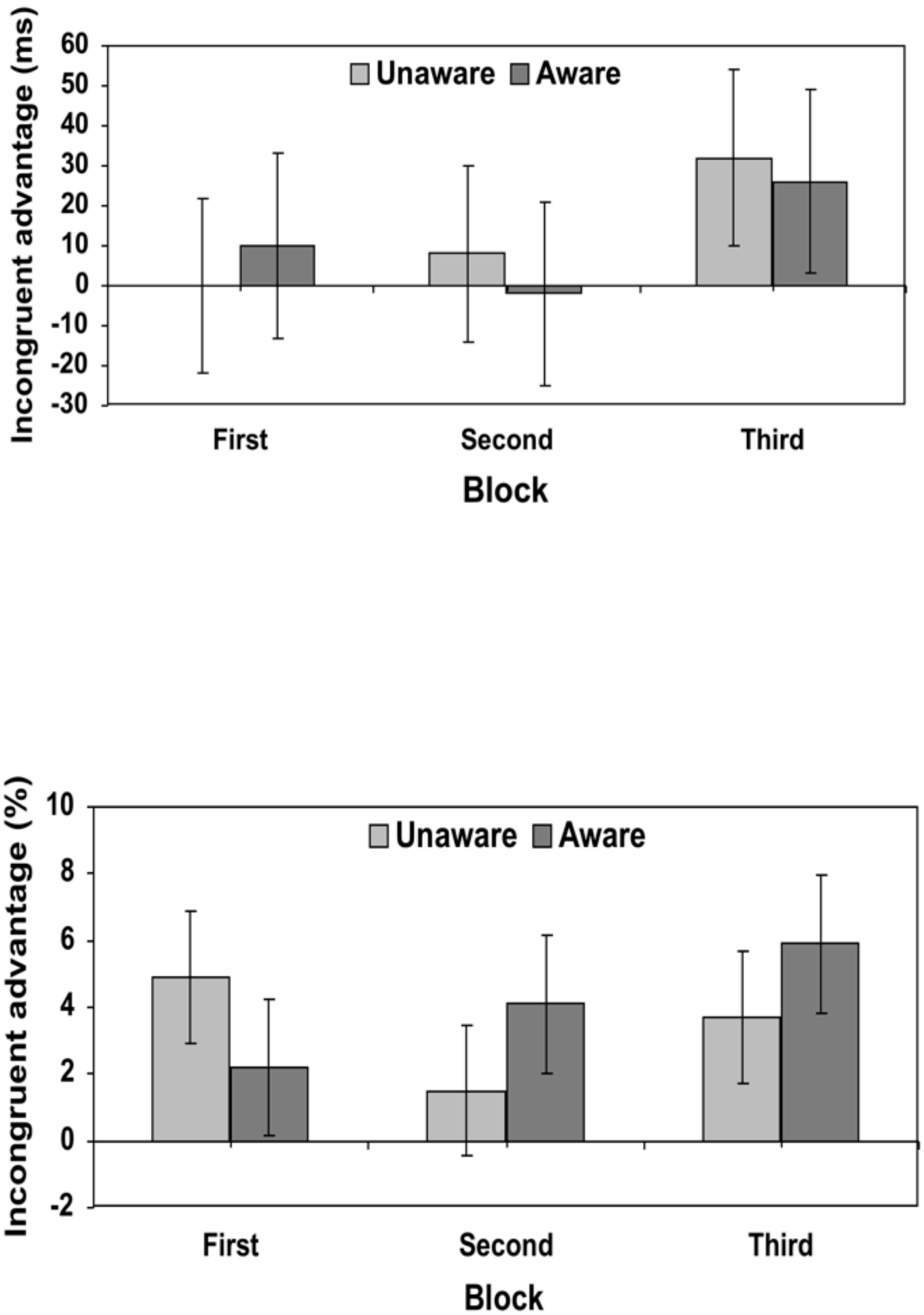

Figure 1. Reverse congruence effect as a function of Block and Proportion awareness. (Error bars = standard error of the mean.) Top panel = RT, Bottom panel = Error rate.

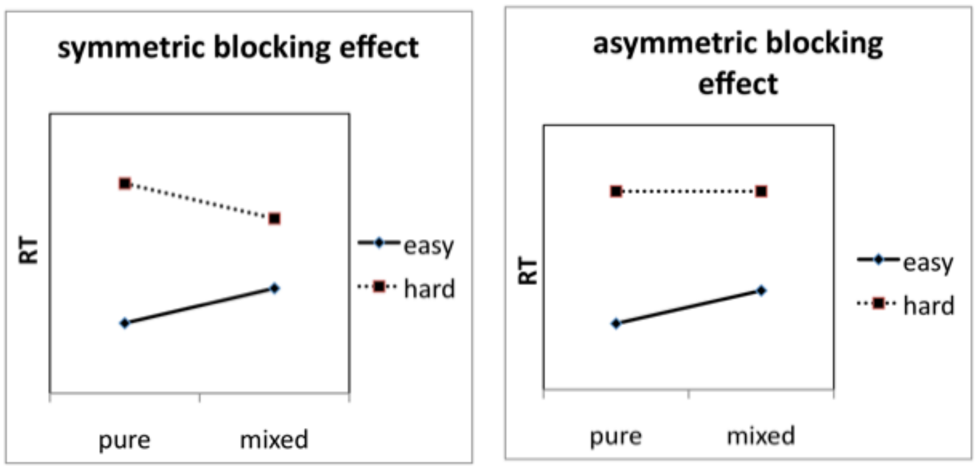

Figure 2. Symmetric and asymmetric blocking effects

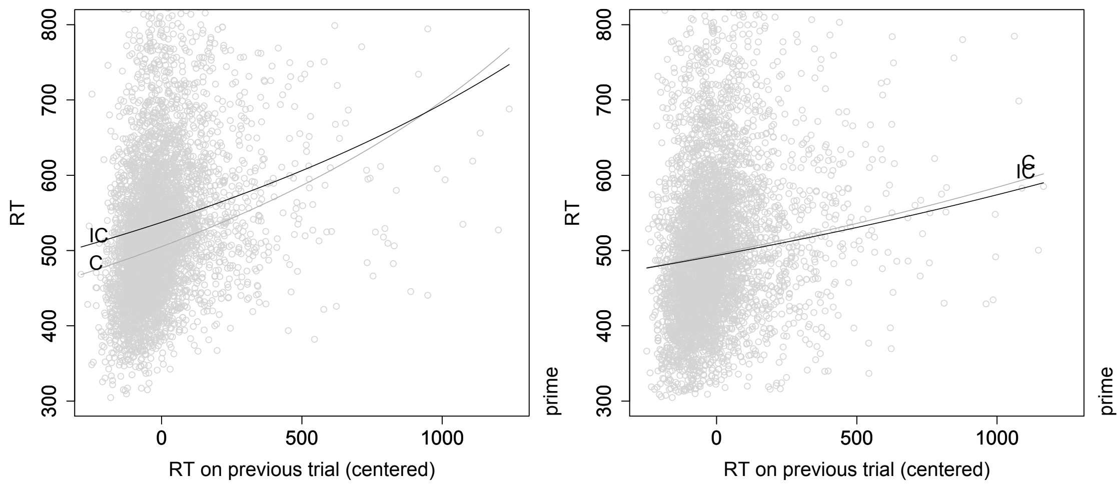

Figure 3 Interaction effects between RT of previous trial (centered) and Congruence in Experiment 2. Left panel = Masked prime group, right panel = Visible prime group. Each circle represents an observation.

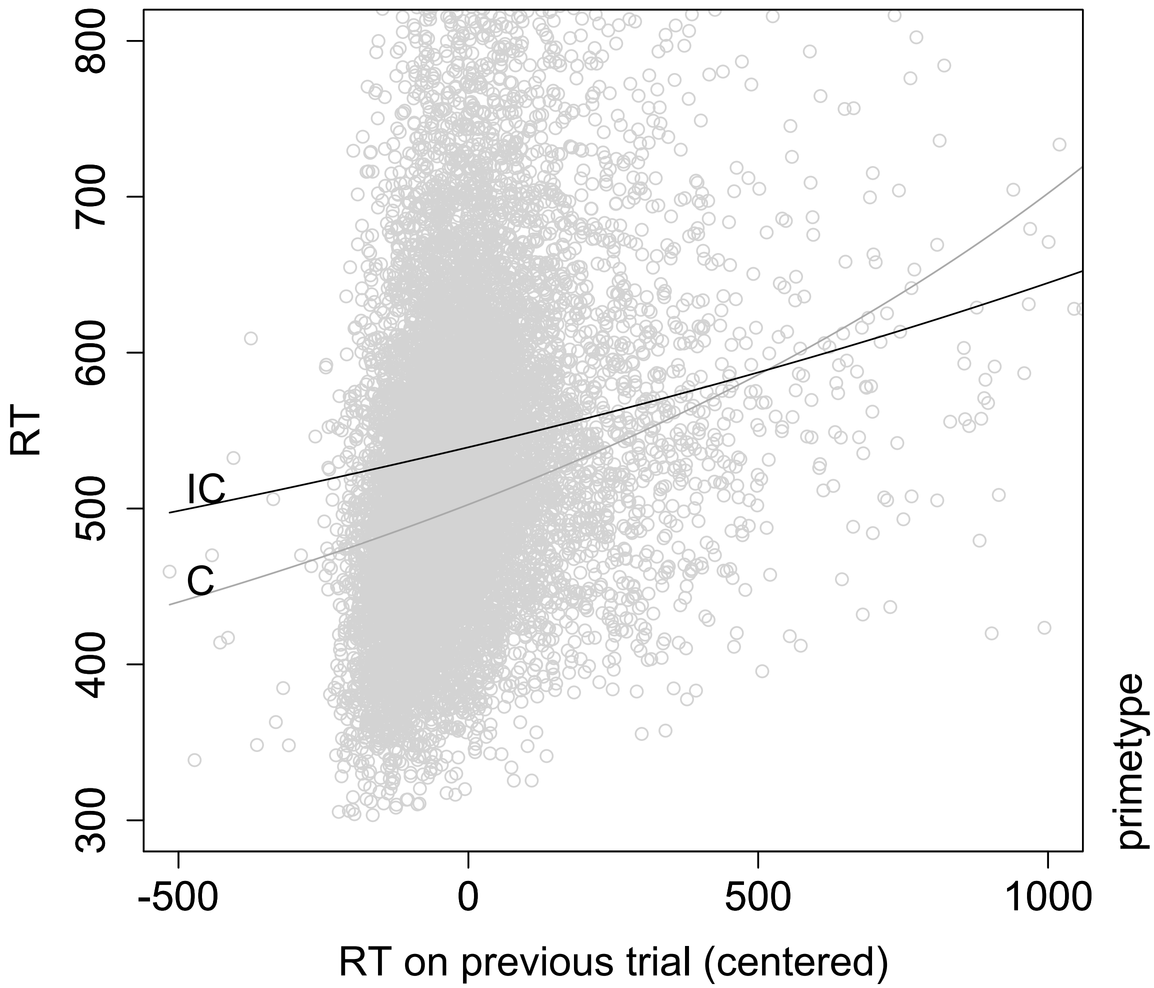

Figure 4. Interaction effects between RT of previous trial (centered) and Congruence in Experiment 3. (C = congruent, IC = incongruent). Each circle represents an observation.

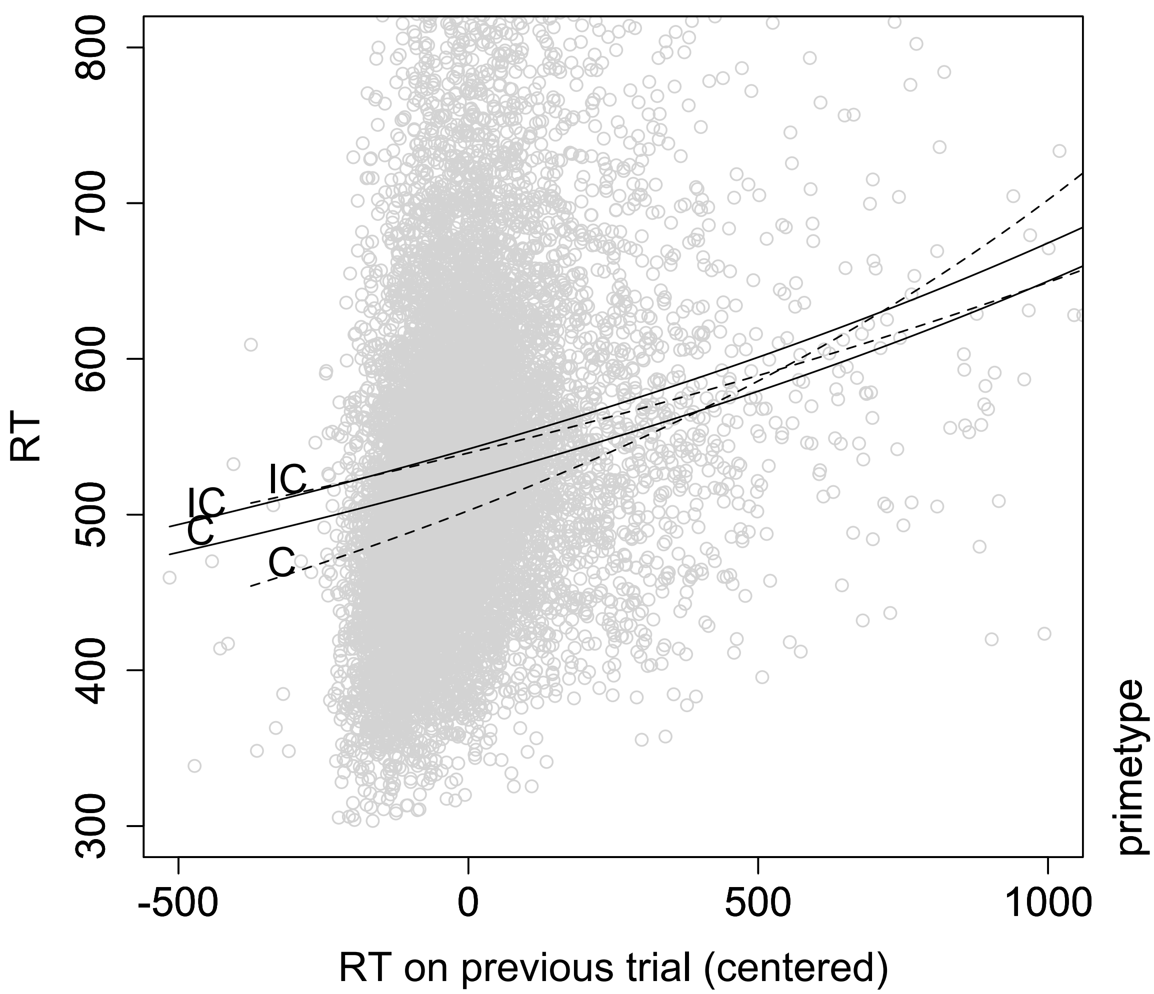

Figure 5. Three-way interaction effect between RT of previous trial (centered), Congruence and Proportion in Experiment 3. (Dashed lines = high proportion congruent block; solid lines = low-proportion congruent block; C = congruent, IC = incongruent). Each circle represents an observation.