Perception & Psychophysics
2008, 70 (4), 680-687
doi: 10.3758/PP.70.4.680

Contingent attentional capture occurs by activated target congruence

Atsunori Ariga and Kazuhiko Yokosawa
University of Tokyo, Tokyo, Japan

Contingent attentional capture occurs when a stimulus property captures an observer's attention, usually related to the observer's top-down attentional set for target-defining properties. In this study, we examined whether contingent attentional capture occurs for a distractor that does not share the target-defining property at a physical level, but does share that property at an abstract level of representation. In a rapid serial visual presentation stream, we defined the target by color (e.g., a green-colored Japanese kanji character). Before the target onset, we presented a distractor that referred to the target-defining color (e.g., a white-colored character meaning "green"). We observed contingent attentional capture by the distractor, which was reflected by a deficit in identifying the subsequent target. This result suggests that because of the attentional set, stimuli were scanned on the basis of the target-defining property at an abstract semantic level of representation.

Attentional selection of information from a stimulus array is controlled in at least two distinct ways (Yantis, 1993). One way involves the viewer's ability to control what regions or objects should be selected on the basis of the viewer's goals or intentions (i.e., *goal-directed*, or *top-down*, *selection*). The other way is *stimulus-driven*, or *bottom-up*, *selection*, which refers to the fact that certain properties of a stimulus may capture attention independently of the viewer's current goals or intentions. Whereas goal-directed selection generally helps viewers' performance, stimulus-driven selection potentially disrupts it, because needless selection sometimes occurs. This article focuses on an interesting interaction between the viewers' top-down intentions and stimulus-driven attentional selection, which is known as *contingent attentional capture*. Contingent attentional capture is a phenomenon in which the viewer's attention is driven to select (or is captured by) a stimulus that looks like the current target.

Contingent attentional capture has conventionally been explained in terms of needless shifts of spatial attention. For example, Folk, Remington, and Johnston (1992) found that responses to red targets were slower when they were preceded by a nonpredictive, or to-be-ignored, red cue that was presented at a spatial location different from the target location. They argued that because the target-defining property was the color red, all red objects were capable of capturing attention. Thus, when a red cue was presented at the wrong location, attention was involuntarily shifted to that location, which was a needless shift because an additional spatial shift was required to redirect attention to the target location. As alternative evidence of contingent attentional capture, Folk, Leber, and Egeth (2002) showed a transient deficit in target processing in a rapid serial visual presentation (RSVP) task. Observers viewed an RSVP stream of letters and searched for a target letter defined by a specific color (e.g., green). Before the target appearance, four distractors (#s) were presented at spatial locations adjacent to the RSVP stream. Identification accuracy for the target dropped sharply when one of the #s was presented in the target-defining color (e.g., green), whereas accuracy stayed constant when the distractors were a different color (e.g., gray). The obtained result was interpreted as a transient deficit in target processing caused by contingent attentional capture—that is, the needless shift of spatial attention away from the RSVP stream.

In summary, it is likely that the selection of targets requires observers to establish a top-down attentional set for the target-defining property (e.g., the color green). Then, using the established attentional set, the observers select the corresponding stimuli (e.g., green objects) as potential targets, whether or not they are really targets, on the basis of bottom-up stimulus properties processed preattentively. Importantly, in order to make target selection efficient, the process of contingent attentional capture is automatic and obligatory, although this does involve some cost. Folk and colleagues' primary conclusion was that a given stimulus property will contingently capture observers' attention only to the extent that it matches the top-down attentional set established by observers to efficiently perform the task at hand.

Although there is considerable evidence for contingent attentional capture mediated by spatial shifts of attention, recent research has shown that a similar deficit in target processing can also be obtained under conditions in which spatial factors are ruled out (see, e.g., Barnard, Scott, Taylor, May, & Knightley, 2004; Folk, Leber, &

A. Ariga, ariga@L.u-tokyo.ac.jp

Egeth, in press; Ghorashi, Zuvic, Visser, & Di Lollo, 2003; Maki & Mebane, 2006; Visser, Bischof, & Di Lollo, 2004). For example, Barnard et al. demonstrated contingent attentional capture within a single RSVP stream of words by manipulating the words' meanings. They presented an RSVP stream of filler words that referred to natural things (e.g., *cloud* and *rainbow*) and required observers to report an inserted target word that referred to people who are paid (e.g., *waitress* and *banker*). Prior to the target onset, a critical distractor word was presented. They found that how often observers missed the target depended on whether the critical distractor was semantically related to the target. Closely related distractors (e.g., referring to people who are not paid, such as *shopper* and *tourist*) more frequently disrupted observers' target identifications than did unrelated distractors (e.g., household items, such as *freezer* and *television*). The deficit in target processing largely depended on the amount of attentional capture—that is, the extent to which the distractor matched the observers' attentional set, established to efficiently process the target.

Barnard et al.'s (2004) finding of a deficit in a single RSVP stream can also be discussed in terms of the *attentional blink* (AB), which is reflected by an impairment in identifying the second of two targets presented with a short temporal lag (Raymond, Shapiro, & Arnell, 1992). The precise cause of this second-target deficit is still undetermined, but there is general agreement that it stems from a transient delay in the allocation of attention to the second target because of the selection of the first (see, e.g., Chun & Potter, 1995; Shapiro, Arnell, & Raymond, 1997). Thus, Barnard et al.'s target-like distractor may have corresponded to the first target in the AB paradigm. Since contingent attentional capture is the automatic and obligatory selection of items that match observers' attentional set, the needless selection of a target-like distractor may impair the processing of the following target because of the same mechanisms underlying the AB (Ghorashi et al., 2003; Spalek, Falcon, & Di Lollo, 2006; Visser et al., 2004).

Nonspatial attentional selection, like spatial selection, is contingent on the top-down attentional set. In a recent study, Folk et al. (in press) proposed that contingent attentional capture in the time domain shares several characteristics with that in the spatial domain. Although it is unclear whether a common mechanism underlies both spatial and nonspatial contingent attentional capture, further investigations "are clearly needed to explore the precise conditions under which nonspatial attention capture emerges" (Folk et al., in press). This topic is the focus of this article.

Three things are notable about the function of the attentional set that leads to contingent attentional capture, irrespective of whether it is spatial or nonspatial. First, the attentional set established by an observer is used to scan stimuli on the basis of the target-defining property. In several experiments, contingent attentional capture occurred for a distractor that shared the target-defining (or a similar) property (e.g., Bacon & Egeth, 1994; Barnard et al., 2004; Folk et al., 2002, in press; Folk et al., 1992; Spalek et al., 2006). Second, as reported by Barnard et al., the attentional set focuses on several properties (including the meaning) of stimuli. This means that contingent attentional capture occurs not only in relation to the perceptual properties of stimuli (e.g., color; Folk et al., 2002, in press; Folk et al., 1992; Spalek et al., 2006), but also for higher representations activated by stimuli. Third, the degree of attentional capture by a distractor depends profoundly on the features it shares with the target (Barnard et al., 2004; Maki & Mebane, 2006; Spalek et al., 2006). A distractor that substantially matches the observers' attentional set may capture attention considerably, and also may elicit a greater cost in performance.

These findings suggest that the content of the attentional set indeed establishes the target-defining property, and that attentional selection based on this set is the important step in scanning to determine whether an incoming stimulus is a target or a distractor. This gives rise to the question of the representation level at which the target-defining property is scanned by the attentional set. Previous studies have suggested that the target-defining property is scanned at the physical level. If this property is the color green, then the content of the attentional set is also the color green. In this case, the attentional set could focus only on the color dimension of stimuli (e.g., Folk et al., 2002, in press). We propose the new idea that the target-defining property is scanned according to the attentional set at an abstract level of representation. For example, if the target-defining property is the visual color green, the content of the attentional set might include abstract representations of green (for example, the color green as well as the meanings of green, such as those represented by different linguistic characters or symbols for green). In this case, incoming stimuli are scanned on the basis of several dimensions (e.g., the color and meaning dimensions). We examined this idea in our experiments.

There is a long history of studying Stroop effects (MacLeod, 1991; Stroop, 1935), in which response systems are captured by semantic-level confusions between two dimensions of the stimuli (e.g., the color and meaning of a word). In other words, the visual system has a temporary inability to distinguish two different properties of the same stimulus. Virzi and Egeth (1984) demonstrated clear evidence for such confusion across dimensions. They presented several words simultaneously in a display, each of which was printed in a different color (e.g., BIG in red ink and BLUE in brown ink). Observers were required to report as many words as they could, along with the colors in which they were printed. The results reflected confusion between the color and meaning dimensions (i.e., between the color of a word and the color name). For instance, in the example above, the observers frequently reported that BIG was in blue ink. Virzi and Egeth argued that the color and meaning properties of stimuli are similarly represented at an abstract semantic level of representation, and hence are confusable, which supports our proposal.

In this study, we presented an RSVP stream of Japanese kanji characters. We defined a target in the color dimension and manipulated a distractor preceding the target with respect to its meaning in order to manipulate the distractor–target congruence across dimensions, as in the Stroop literature (MacLeod, 1991). We use the terms *congruence* and *incongruence* on the basis of whether or not the distractor has the target-defining property at an abstract level of representation. For example, if the target-defining property is the color green, the character that means green would be the congruent distractor, and the character that means red would be an incongruent distractor. We hypothesized that if the target-defining property is scanned according to the attentional set at an abstract level of representation, nonspatial contingent attentional capture by the congruent distractor would occur, and consequently, the correct identification of the target would be reduced—that is, the AB would occur.

GENERAL METHOD

Apparatus

Stimuli were displayed on an EIZO FlexScan E67T monitor controlled by a PC/AT-compatible computer equipped with a Cambridge Research Systems VSG 2/5 frame store. Observers viewed the display from a distance of about 57 cm in a dark room.

Stimuli

The stimuli were a small cross fixation point that subtended 0.2º × 0.2º of visual angle and single Japanese kanji characters that subtended 0.6º × 0.6º. As in Barnard et al. (2004), 22 characters referred to natural things (e.g., 谷 "valley" and 星 "star"), and 2 referred to color names (緑 "green" and 赤 "red"); thus, a total of 24 characters were prepared (see the Appendix). The mean familiarity and complexity of the natural-thing characters were nearly equal to those of the color-name characters (according to the NTT communication database; Amano & Kondo, 1999). The natural-thing characters neither induced specific color images nor shared the same forms or pronunciations with each other or with the color-name characters. The fixation point and the characters were displayed in white (47.9 cd/m^2, CIE [.263, .282]) on a black (0 cd/m^2) background, except for the target character, which appeared either in green (30.0 cd/m^2, CIE [.299, .581]), in Experiments 1A and 2A, or in red (22.1 cd/m^2, CIE [.603, .343]), in Experiments 1B and 2B.

Procedure

Observers started each trial by pressing the space key, and the fixation point was then presented at the center of the screen, where the kanji characters would subsequently appear. After a delay of 500 msec, the RSVP of temporal frames began. Each frame was a presentation of a character for either 120 msec (Experiments 1A and 1B) or 100 msec (Experiments 2A and 2B). The presentation was continuous, with no blank screen between images (interstimulus interval = 0 msec).

In each RSVP sequence, 22 characters were presented. Of the 22 characters, the 20 fillers and the 1 colored target were natural-thing characters; the other character was the distractor. Each natural-thing character was presented equally often throughout the experiment and was the target equally often. The distractor character, presented within the 11th to the 14th frame and always preceding the target appearance, was manipulated as a factor of distractor type, in *congruent*, *incongruent*, and *neutral* conditions. In the congruent and incongruent conditions, the distractors were color-name characters (meaning "green" or "red"). In the congruent condition, the meaning of the distractor was congruent with the target-defining color (Figure 1A). In the incongruent condition, the meaning of the distractor was incongruent with the target-defining color (Figure 1B). In the neutral condition, the distractor was a natural-thing character, similar to the fillers (Figure 1C). In addition, the temporal interval between the onsets of the distractor and the target was manipulated as a factor of distractor–target (D–T) lag: 1, 2, 3, and 7. Each temporal lag corresponded to a distractor–target stimulus onset asynchrony (SOA) of 120, 240, 360, or 840 msec, respectively, in Experiments 1A and 1B, or of 100, 200, 300, and 700 msec, respectively, in Experiments 2A and 2B. *Lag 1* means that the target immediately followed the distractor, whereas the later lags reflect the postdistractor serial position of the target. The observers were given 24 practice trials prior to the experiment, which consisted of a total of 300 trials.

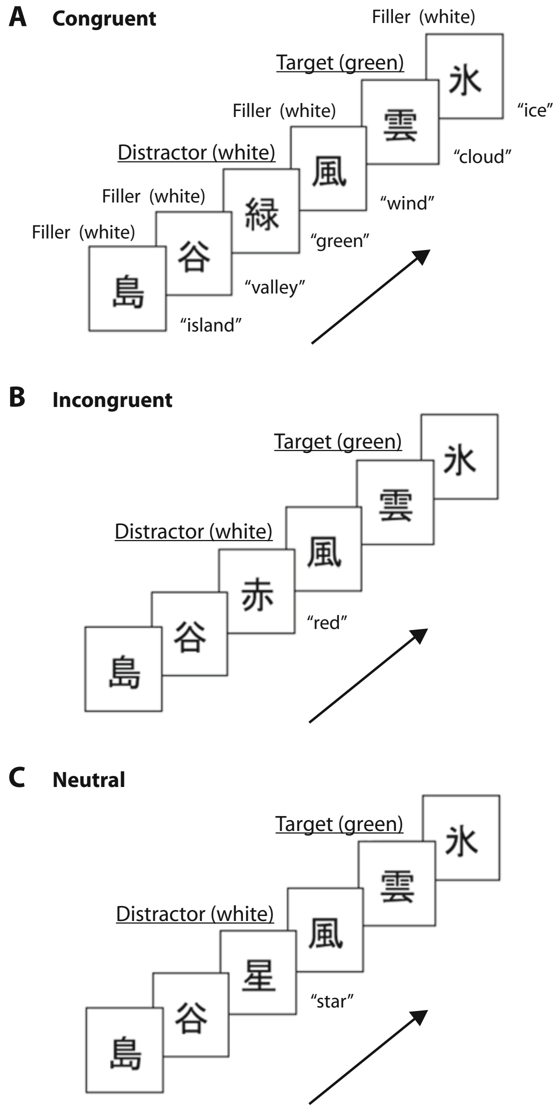

Figure 1. Examples of the trial events when the target is green and the D–T lag is 2. This figure shows each distractor-type condition: congruent (A), incongruent (B), and neutral (C).

Design

A 3 (distractor type: congruent, incongruent, neutral) × 4 (D–T lag: 1, 2, 3, 7) within-observers design was used.

Task

The observers' task was to identify the target character. After the presentation of the RSVP stream, a list of eight candidates, which consisted of the target and seven of the fillers presented in the present trial, was presented in the display. The color-name distractors never appeared in the list. The eight candidates were aligned in an imaginary square, and the observers chose the target by pressing the corresponding arrow key on the numeric keypad (e.g., upper right, left below).

EXPERIMENTS 1A AND 1B

In Experiments 1A and 1B, we examined whether nonspatial contingent attentional capture can occur for distractors that do not share the target-defining property at a physical level, but that do at an abstract level of representation. In this study, the target-defining dimension was color. Note that we manipulated the preceding distractor with respect to its meaning. If the occurrence of contingent attentional capture depends on the distractor–target congruence across dimensions, we would obtain an AB; that is, correct target identifications would be reduced at a short D–T lag and would recover as the D–T lag becomes longer.

Method

Observers. Nineteen naive observers (13 males and 6 females) participated in Experiment 1A, and 17 naive observers (11 males and 6 females) participated in Experiment 1B. All were Japanese students ranging from 19 to 26 years old. All had normal or corrected-to-normal visual acuity and normal color vision, according to self-report.

Procedure. The target was defined by its green color in Experiment 1A and its red color in Experiment 1B. The RSVP sequence was presented with a 120-msec SOA in both experiments. In the congruent condition, the distractor was the character for the name of the target-defining color—that is, "green" in Experiment 1A and "red" in Experiment 1B. In the incongruent condition, the distractor character was the other color name—that is, "red" in Experiment 1A and "green" in Experiment 1B.

Results

Experiment 1A. The mean percentage of correct identifications of the target was calculated for each condition separately and is shown in Figure 2A. A 3 (distractor type: congruent, incongruent, neutral) × 4 (D–T lag: 1, 2, 3, 7) two-way ANOVA revealed a significant main effect of D–T lag [$F(3,54) = 5.78$, $p < .005$] but no significant main effect of distractor type [$F(2,36) = 3.09$, n.s.]. The interaction between these factors was significant [$F(6,108) = 3.38$, $p < .005$]. A significant simple main effect of D–T lag was observed only for the congruent condition [$F(3,54) = 12.70$, $p < .001$], and a significant effect of distractor type was observed in the lag-2 condition [$F(2,36) = 5.52$, $p < .01$]. Multiple comparisons (Tukey's HSD test) revealed that, in the congruent condition, correct identifications for the lag-2 condition were significantly lower than those at any other lag ($p < .05$). Furthermore, in the lag-2 condition, correct identifications in the congruent condition were lower than those in the incongruent and neutral conditions ($p < .05$).

Experiment 1B. These results are shown in Figure 2B. The ANOVA revealed a significant main effect of D–T lag [$F(3,48) = 3.79$, $p < .05$], no significant main effect of distractor type [$F(2,32) = 3.20$, n.s.], and no interaction between these factors [$F(6,96) = 1.49$, n.s.].

Discussion

The observers in Experiment 1A tended to miss the target when the D–T lag was short (i.e., the AB occurred) only in the congruent condition. High performance in the lag-1 condition is referred to in studies of the AB as *lag-1 sparing* (Visser, Bischof, & Di Lollo, 1999), which we will discuss in the General Discussion. However, the observers in Experiment 1B showed no such effect; their target identifications were not affected by the distractor type.

We propose that two accounts could explain these asymmetric results: a *contingent-capture* account and a *bottom-up saliency* account. In both, the difference in results between the two experiments would be attributed to a difference in perceptual complexity of the color-name distractors. It is noteworthy that the congruent distractor in Experiment 1A was perceptually a high-complexity stimulus (緑 "green"), as compared with that in Experiment 1B (赤 "red"). According to the contingent-capture account, as discussed in the introduction, as we predicted the target-defining property was basically scanned in line with the attentional set at an abstract level of representation. Even the congruent distractor in Experiment 1B contingently captured the observers' attention to some degree. However, the period of capture in Experiment 1B would have been shorter than for the high-complexity congruent distractor in Experiment 1A, because the low-complexity stimulus could be processed more easily than the high-complexity one. As a consequence, in Experiment 1B the observers' attention was quickly released from the distractor in the congruent condition, and the following target was efficiently processed even with the short D–T lag. The plausibility of this prediction is supported by the finding in AB studies that a first target that is processed easily eliminates the AB (e.g., Chun & Potter, 1995; Raymond et al., 1992; Seiffert & Di Lollo, 1997).

On the other hand, in terms of the bottom-up saliency account, we noticed that the high-complexity distractor was more salient in the RSVP stream, as compared with the low-complexity stimulus, because of a difference in the overall luminance between those stimuli. The term *saliency* is used here only in reference to local stimulus-driven contrast. Because a salient distractor automatically captures observers' attention only in a bottom-up manner (see, e.g., Theeuwes, 1992, 1995, 2004), we could also explain the difference in the results by assuming that the high-saliency distractor had enough power to capture attention and induce the AB independently of observers' top-down attentional set, although the low-saliency distractor did not. However, in this scenario, we should have observed low accuracy as well in the incongruent condition of Experiment 1B (in which the highly salient 緑 "green" was the distractor), but this was not found. Therefore, we consider this account unlikely.

To evaluate these accounts, we manipulated the rate of the RSVP in the next experiments. If the contingent-capture account is correct, shortening the SOA would make the effect of capture visible even with the low-complexity congruent distractor, because the target would appear within the short period of capture. That is, the AB would occur with both target colors only in the congruent condition,

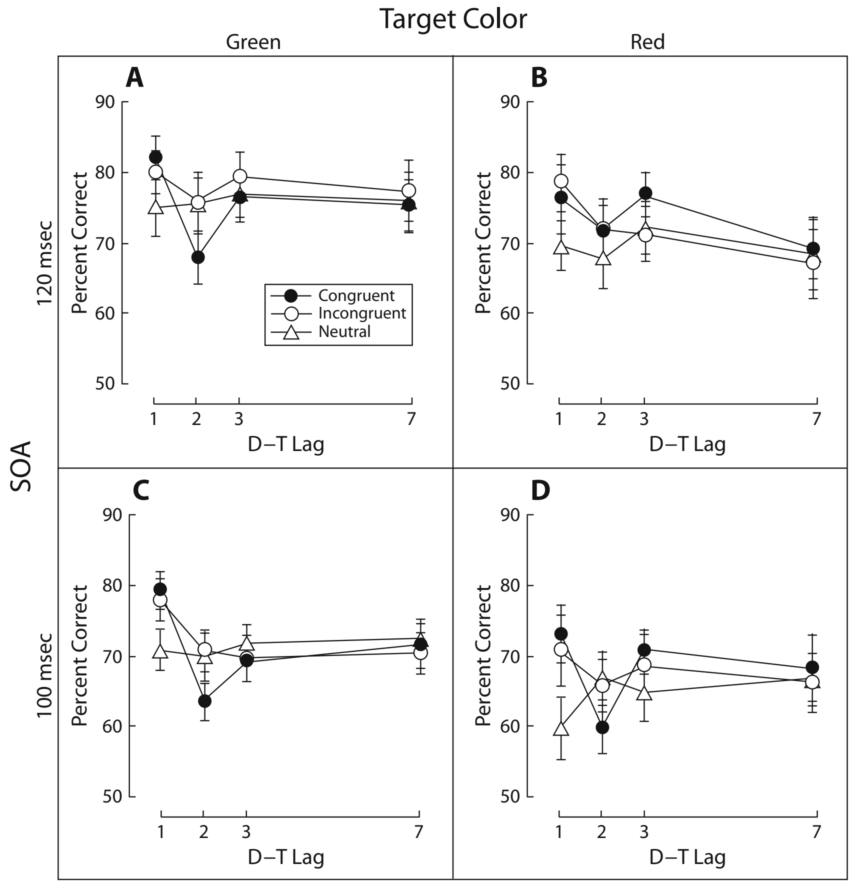

Figure 2. Mean percentages of correct identifications of the target, presented separately for each distractor type and D–T lag condition in Experiments 1A (A), 1B (B), 2A (C), and 2B (D). The bars in each figure indicate the standard errors of the means.

as in Experiment 1A. However, if the bottom-up saliency account is correct, the effect of capture (or the AB) would not occur with the low-complexity congruent distractor, because the local stimulus complexity would be little affected by the rate of the RSVP. Instead, we might observe the AB with the high-complexity incongruent distractor.

EXPERIMENTS 2A AND 2B

Method

Observers. Nineteen naive observers (12 males and 7 females) participated in Experiment 2A, and 21 naive observers (16 males and 5 females) participated in Experiment 2B. All were Japanese students ranging in age from 19 to 28 years. All had normal or corrected-to-normal visual acuity and normal color vision, according to self-report.

Procedure. The target was defined by its green color in Experiment 2A and its red color in Experiment 2B. The RSVP sequence was presented with a 100-msec SOA in both experiments.

Results

Experiment 2A. The results are shown in Figure 2C. The ANOVA revealed a significant main effect of D–T lag [$F(3,54) = 7.29$, $p < .001$] but no significant main effect of distractor type [$F(2,36) = 0.56$, n.s.]. The interaction between these factors was significant [$F(6,108) = 3.13$, $p < .01$]. As in Experiment 1A, a significant simple main effect of D–T lag was observed in the congruent condition [$F(3,54) = 12.37$, $p < .001$], and a significant effect of distractor type was observed in the lag-2 condition [$F(2,36) = 4.85$, $p < .05$]. Multiple comparisons revealed that in the congruent condition, correct identifications for the lag-2 condition were significantly lower than those for the lag-1 and lag-7 conditions ($p < .05$). Furthermore, in the lag-2 condition, correct identifications in the congruent condition were lower than those in the incongruent and neutral conditions ($p < .05$).

Experiment 2B. The results are shown in Figure 2D. The ANOVA revealed a significant main effect of distractor type [$F(2,40) = 4.25$, $p < .05$] and no significant main effect of D–T lag [$F(3,60) = 2.01$, n.s.]. The interaction between these factors was significant [$F(6,120) = 5.29$, $p < .001$]. As in Experiments 1A and 2A, a significant simple main effect of D–T lag was observed in the congruent condition [$F(3,60) = 5.69$, $p < .005$], and a significant effect of distractor type was observed in the lag-2 condition [$F(2,40) = 5.34, p < .01$]. Multiple comparisons revealed that in the congruent condition, correct identifications for the lag-2 condition were significantly lower than those at any other lag ($p < .05$). Furthermore, in the lag-2 condition, correct identifications in the congruent condition were lower than those in the incongruent and neutral conditions ($p < .05$).

Discussion

The expected results were obtained in both experiments. Deficits in target identifications for the short D–T lag were observed only when the meaning of the distractor was congruent with the target-defining color, supporting the contingent-capture account. Therefore, we can conclude that nonspatial contingent attentional capture by the distractor occurred for distractors that shared the target-defining color at an abstract level of representation, even without the physical target-defining color.

GENERAL DISCUSSION

In this article, we have proposed a new idea, that contingent attentional capture by the distractor can occur for a distractor that does not share the target-defining property at a physical level but that does share that property at an abstract level of representation. In all of these experiments, we defined the target by color in an RSVP stream and manipulated congruence across dimensions (i.e., the target-defining color and the meaning of the distractor), as in the Stroop literature (MacLeod, 1991). The results were clear: Observers more frequently missed the target when the meaning activated by the distractor was congruent with the target-defining color than when the distractor was incongruent or neutral. The AB occurred only in the congruent condition, which was evidence of contingent attentional capture in the time domain. In this study, we demonstrated that the visual system could be confused between the meaning of a distractor and the color of the target. The color and meaning properties would both be represented at an abstract semantic level of representation, as suggested by Virzi and Egeth (1984). Therefore, nonspatial contingent attentional capture occurred with a distractor that had the target-defining property only at an abstract level of representation.

There is generally accepted evidence that a singleton in the visual field captures observers' spatial attention. For example, Theeuwes (1992, 2004) used a visual search task (known as the *additional singleton task*) in which two singletons were presented simultaneously, one of which (a shape singleton) was the target and the other (an additional color singleton) the critical distractor. Observers' responses to the target were slower when the critical distractor was present than when it was absent. Theeuwes's (1992, 2004) claim was that the additional singleton automatically captured observers' attention in a bottom-up manner and disrupted performance. Considering that the congruent and incongruent distractors in this study were both singletons in the meaning dimension (color names among natural things), according to Theeuwes's (1992, 2004) claim, both the congruent and incongruent distractors should result in deficits in performance. This was not the case, however; only the congruent distractor elicited a significant deficit. Therefore, the present phenomenon does not reflect the bottom-up component of attentional capture, but is instead an effect of observers' top-down attentional set on attentional selection in the time domain.

One may argue that the effect we have obtained looks somewhat smaller than the usual AB deficit that occurs at lags 2–5 (see, e.g., Chun & Potter, 1995; Raymond et al., 1992). We consider that there are two possible explanations for this, as follows. First, the present task was to report only one target while ignoring the distractor, whereas the task in conventional AB studies has been to report two targets. This suggests that the visual system was likely to be disengaged from our to-be-ignored distractor earlier than from a to-be-reported first target, and each of them consequently impaired later processing of the target accordingly. In fact, the present results are consistent with the findings of Chun (1997), who reported a similar small AB deficit at lags 1 and 2 in a condition in which observers were required to ignore the first target. Second, it has been recently reported that the magnitude of the AB deficit is affected by the similarity between the to-be-ignored first target and the to-be-reported second target; that is, the AB deficit is greater with higher similarity (Spalek et al., 2006). Because the distractor in our study did not share the physical target-defining color but did share the target-defining color at an abstract level of representation, distractor–target similarity would have been low and should have produced only a small AB deficit. Considering these two explanations, it is plausible that the effect we obtained would look somewhat small. However, it is noteworthy that similar effects were robustly observed in three of our four experiments, independently of target color.

We can adequately explain the present findings with the temporary loss of control (TLC) model proposed by Di Lollo, Kawahara, Ghorashi, and Enns (2005). The TLC model explains the AB by focusing on the state of the attentional set. In this model, items are identified through two consecutive stages; items selected by the attentional set in the first stage enjoy further attentional processing. Importantly, the TLC model predicts that the attentional set is disrupted by nontarget items when attentional processing is busy because the attentional set and attentional processing are controlled by a common system. Thus, when nontarget items follow the first target, they disrupt the attentional set, and the second target cannot be selected; as a result, the AB occurs. On the other hand, when no nontarget items occur between the two targets (i.e., the second target immediately follows the first), the second target can be selected by the intact attentional set, which is called *lag-1 sparing* (Visser et al., 1999). In our study, the congruent distractor was selected by the attentional set as though it was the target because the activated representation was congruent with the target-defining property. Then, a filler item following the distractor disrupted the attentional set, and the following target was often not selected; in other words, the AB occurred. However, when the target immediately followed the congruent distractor, the target was efficiently selected by the intact attentional set, in much the same fashion as in lag-1 sparing. That is why we observed significantly higher performance for the lag-1 condition than for the lag-2 condition with congruent distractors in all experiments except Experiment 1B.

However, correct identifications were also higher at lag 1 in the incongruent condition than in the neutral condition, although a significant effect was obtained only in Experiment 2B. This is most likely because the response display, a list of eight candidates for the target, did not contain the color-name distractors. Our logic is the following: It is known that observers in an RSVP task with Japanese kanji characters often perceive a distractor immediately preceding the target as the target (a phenomenon known as *illusory conjunction*; see, e.g., Kikuchi, 1996). Also, the observers in our study might often confuse the congruent, incongruent, and neutral distractors for the target in the lag-1 condition. However, in the congruent and incongruent conditions, their confusion would be diminished, because the color-name distractors never appeared in the list of the target candidates. On the other hand, in the neutral condition, they could be confused because the neutral distractor (a natural-thing character) could appear in the response display. Thus, observers could more accurately choose the target from the response display at lag 1 in the congruent and incongruent conditions, as compared with the neutral condition.

We believe that our findings are not kanji-specific, but instead reflect a more general phenomenon. The processing of logographic scripts such as kanji seems to have more in common with picture processing than with the processing of syllabic scripts. It has typically been argued that access to meaning for kanji is achieved directly from print without phonological mediation. Thus, kanji possess some direct means to access a lexical code (see, e.g., Goryo, 1987). However, it has been reported that meaning can be activated preattentively, even in an RSVP stream, independently of whether the script is logographic (Kikuchi, 1996) or syllabic (see, e.g., Luck, Vogel, & Shapiro, 1996; Maki, Frigen, & Paulson, 1997). Therefore, it is plausible to predict that, irrespective of script, activated representations could be scanned in line with the attentional set and could elicit nonspatial contingent attentional capture as long as the distractor has the target-defining property at an abstract level of representation; this prediction awaits further investigation.

In summary, our study demonstrates that nonspatial contingent attentional capture can occur with a distractor that is congruent with the target-defining property across dimensions (meaning and color). The results suggest that the color and meaning properties of stimuli are similarly represented at an abstract semantic level of representation, a level at which the visual system might confuse these properties. This study has revealed one condition in which nonspatial attentional capture emerges. It would be interesting to examine whether this phenomenon is also observed in the spatial capture paradigm. We hope that this phenomenon and future investigations related to it will lead to a better understanding of the nature of attentional selection.

AUTHOR NOTE

This research was supported by grants from the Japan Society for the Promotion of Science to A.A. and K.Y. We thank Miyou Ninomiya for collecting data and two reviewers for helpful comments on an earlier version of the manuscript. Correspondence should be addressed to A. Ariga, Department of Psychology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 113-0033, Japan (e-mail: ariga@L.u-tokyo.ac.jp).

REFERENCES

Amano, N., & Kondo, K. (1999). *Nihongo-no goitokusei* [Lexical properties of Japanese] (1st ed.). Tokyo: Sanseido.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, **55**, 485-496.

Barnard, P. J., Scott, S., Taylor, J., May, J., & Knightley, W. (2004). Paying attention to meaning. *Psychological Science*, **15**, 179-186.

Chun, M. M. (1997). Temporal binding errors are redistributed by the attentional blink. *Perception & Psychophysics*, **59**, 1191-1199.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 109-127.

Di Lollo, V., Kawahara, J.-I., Ghorashi, S. M. S., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research*, **69**, 191-200.

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics*, **64**, 741-753.

Folk, C. L., Leber, A. B., & Egeth, H. E. (in press). Top-down control settings and the attentional blink: Evidence for nonspatial contingent capture. *Visual Cognition*.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 1030-1044.

Ghorashi, S. M. S., Zuvic, S. M., Visser, T. A. W., & Di Lollo, V. (2003). Focal distraction: Spatial shifts of attentional focus are not required for contingent capture. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 78-91.

Goryo, K. (1987). *Yomuto iu koto* [Psychology of reading]. Tokyo: Tokyo University Press.

Kikuchi, T. (1996). Detection of kanji words in a rapid serial visual presentation task. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 332-341.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, **383**, 616-618.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.

Maki, W. S., Frigen, K., & Paulson, K. (1997). Associative priming by targets and distractors during rapid serial visual presentation: Does word meaning survive the attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1014-1034.

Maki, W. S., & Mebane, M. W. (2006). Attentional capture triggers an attentional blink. *Psychonomic Bulletin & Review*, **13**, 125-131.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 849-860.

Seiffert, A. E., & Di Lollo, V. (1997). Low-level masking in the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1061-1073.

Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences*, **1**, 291-296.

Spalek, T. M., Falcon, L. J., & Di Lollo, V. (2006). Attentional blink and attentional capture: Endogenous versus exogenous control over paying attention to two important events in close succession. *Perception & Psychophysics*, **68**, 674-684.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, **51**, 599-606.

Theeuwes, J. (1995). Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics*, **57**, 637-644.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review*, **11**, 65-70.

Virzi, R. A., & Egeth, H. E. (1984). Is meaning implicated in illusory conjunctions? *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 573-580.

Visser, T. A. W., Bischof, W. F., & Di Lollo, V. (1999). Attentional switching in spatial and nonspatial domains: Evidence from the attentional blink. *Psychological Bulletin*, **125**, 458-469.

Visser, T. A. W., Bischof, W. F., & Di Lollo, V. (2004). Rapid serial visual distraction: Task-irrelevant items can produce an attentional blink. *Perception & Psychophysics*, **66**, 1418-1432.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 676-681.

APPENDIX
Japanese Single-Kanji Characters Used in This Study (and Their Meanings)

Natural Things				
林 (grove)	幹 (trunk)	星 (star)	原 (field)	藪 (bush)
虹 (rainbow)	峠 (mountain pass)	谷 (valley)	砂 (sand)	島 (island)
野 (field)	潮 (tide)	氷 (ice)	雲 (cloud)	風 (wind)
園 (garden)	丘 (hill)	畔 (dike)	岩 (rock)	泉 (waterhole)
陸 (land)	畑 (plantation)			
Color Names				
	緑 (green)		赤 (red)	

(Manuscript received September 5, 2006;
revision accepted for publication December 17, 2007.)