Contingent Capture at a very short SOA: Evidence against Rapid Disengagement

Online Publication Date: 01 August 2007

To cite this Article: Chen, Peggy and Mordkoff, J. Toby (2007) 'Contingent Capture at a very short SOA: Evidence against Rapid Disengagement', Visual Cognition, 15:6, 637 - 646

To link to this article: DOI: 10.1080/13506280701317968

URL: http://dx.doi.org/10.1080/13506280701317968
Contingent Capture at a very short SOA: Evidence against Rapid Disengagement

Peggy Chen and J. Toby Mordkoff

Pennsylvania State University, University Park, PA, USA

Salient cues that do not contain the target’s defining attribute do not cause a spatial cueing effect in response time when presented 150 ms before the target. This can be explained by Contingent Capture (under which incongruent cues do not capture attention) and also by Rapid Disengagement (under which salient but incongruent cues cause only brief capture). To discriminate between these models, a very short SOA of 35 ms was employed in the standard paradigm. As is true for longer SOAs, there was no evidence of attentional capture by salient but incongruent cues. Thus, the evidence to date favours Contingent Capture over Rapid Disengagement.

It is quite clear that visual attention can be involuntarily captured by uninformative events under certain conditions (e.g., Jonides, 1981; Jonides & Yantis, 1988; Posner & Cohen, 1984). It is less clear whether attentional capture is purely stimulus-driven or depends on top-down control settings in order to occur. The strongest evidence in favour of top-down control comes from work using the Cue Congruence Paradigm (aka the Contingent Capture Paradigm; for an introduction, see Folk & Remington, 2006). In these experiments, the defining attribute of the target is manipulated (between participants or across blocks) and the ability of various, exogenous cues to affect performance is assessed. In some blocks, for example, the participants might be asked to respond to a singleton, onset target (see Panel A of Figure 1); in other blocks, the target might be defined as the one item that is purple, with the others being green (see Panel B). At the same...
Figure 1. Display events in each of the four main conditions with examples of valid and invalid cues. The background was black. The fixation point and placeholder boxes were grey. Other items shown in black were white, those in grey were green, and hollow or dashed items were purple.
time, various cues are presented prior to the target, and these can also be an onset or purple among greens.

The standard finding from this paradigm is a three-way interaction between the type of cue, the defining attribute of the target, and whether the cue and target appear in the same or different locations. When the cue matches the defining attribute of the target (e.g., an onset cue when the target is also onset), a spatial cuing effect is observed: Responses are faster when the cue and target are presented in the same location. Conversely, when the cue does not match the target (e.g., an onset cue when the target is purple), it does not matter if the cue and target are in the same or different locations. In other words: A spatial validity effect is observed with congruent cues, but not with incongruent cues (see, e.g., figure 2 in Folk, Remington, & Johnston, 1992, for example data).

Based on these and other findings, Folk and colleagues have argued in favour of Contingent Capture (e.g., Folk & Annett, 1994; Folk & Remington, 1998, 2006; Folk, Remington, & Johnston, 1992; Folk, Remington, & Wright, 1994). This model asserts that attentional capture depends upon top-down control settings. The task—most notably the definition of the target—determines the attribute for which attention is set. Any stimulus containing this attribute, including uninformative cues, will capture attention; no stimulus that does not contain this attribute will capture attention, regardless of bottom-up salience. Thus, congruent cues cause a spatial validity effect because congruent cues capture attention. Incongruent cues do not cause a spatial validity effect because incongruent cues do not capture attention.

There is, however, an alternative explanation for the results found using the Cue Congruence Paradigm which centres on the typical, 150-ms SOA between the cue and target. Most notably, Theeuwes and colleagues (e.g., Theeuwes, 1992, 1994, 1996, 2004; Theeuwes, Atchley, & Kramer, 2000) have argued that, regardless of intention or goals (i.e., top-down control settings), highly salient stimuli always capture attention in a purely bottom-up manner. In order to explain why a spatial validity effect is only observed with congruent cues, it is also asserted that the duration of capture depends on the match between the cue and the target’s defining attribute (see, also, Lupiañez, Milliken, Solano, Weaver, & Tipper, 2001). When the cue is congruent, attentional capture by the cue is prolonged, such that a spatial validity effect can be observed at any of various SOAs. When the cue is incongruent, Rapid Disengagement is posited to occur, such that attention is only briefly at the cued location, eliminating the spatial validity effect on incongruent-cue trials with long SOAs.

The clearest way to discriminate between Contingent Capture and Rapid Disengagement is to shorten the delay between the cue and the target. If Contingent Capture is correct, this will not change the pattern of
results: Even at short SOAs, there should be no spatial validity effect for incongruent cues, since incongruent cues are argued to never cause capture. In contrast, if Rapid Disengagement is correct, at short SOAs, a spatial validity effect should be found for any salient cue, regardless of congruence, since this model posits attentional capture (of some duration) when the cue has bottom-up salience.

One might be tempted to use an SOA of zero (i.e., simultaneous onset of cue and target), but this should be avoided. Most of all, this condition could produce artifactual evidence of attentional capture by inducing a filtering cost on trials with the cue and target in different locations (see Kahneman, Treisman, & Burkell, 1983). In general, a simultaneous onset greatly increases the chance of the cue being integrated with the target, as opposed to acting separately (see, e.g., Enns & Di Lollo, 1997; Lleras & Moore, 2003). Since the goal of this work is to understand attentional capture by task-irrelevant and uninformative events, any condition that would cause the cue and target to be combined should be avoided. What is needed, therefore, is an experiment with a very short, but nonzero, SOA.

Several short-SOA experiments using the Cue Congruence Paradigm have been conducted, but all appear insufficient for various reasons. For example, the fourth experiment by Folk and Remington (1998) included a 50-ms SOA condition, but only used incongruent cues, so the failure to find a spatial validity effect was not paired with success using congruent cues. In order to be taken as evidence of Contingent Capture, it is crucial that the spatial validity effect be shown to depend on cue congruence (and not just be absent under certain conditions). Another potential problem with this work is the use of an SOA that exactly matches the duration of the cue, such that the target appeared at the same instant that the cues were removed. Several studies have shown that exogenous cues are better able to capture attention when there is some temporal overlap between cues and targets (e.g., Maruff, Yucel, Danckert, Stuart, & Currie, 1999).

The same criticisms hold for the 50-ms condition of the first experiment by Lamy (2005), but this work actually raises another important point by comparing conditions with mixed versus blocked SOAs. In the former case, a spatial validity effect was observed with incongruent cues; a different pattern was observed with blocked SOAs. From this it was argued that mixing SOAs can force participants to use less specific control settings, such that any odd item (regardless of whether it matches the target) is allowed to capture attention (i.e., “singleton detection mode”; Bacon & Egeth, 1994). Thus, Lamy’s findings are not evidence against Contingent Capture, since the “incongruent” cues were also singletons and the attentional system was probably set to be captured by such.

In summary, what is needed is an experiment that includes both congruent and incongruent cues, and employs a fixed but very short SOA.
The cue and target should also overlap in time, as oppose to the cue being removed when the target is presented. Since this particular set of conditions cannot be found in literature, such an experiment was conducted.

**EXPERIMENT**

The present experiment is a near-exact replication of the classic experiment by Folk et al. (1992), which produced the signature three-way interaction between the type of cue, the defining attribute of the target, and the spatial validity of the cue. The one change is the reduction of the SOA between the cue and target from 150 ms to 35 ms.

**Method**

*Subjects.* Thirty-six naïve individuals (mean age of 24.7 years; 21 women; five left-handers) were recruited and paid $6 for participating. All reported normal or corrected-to-normal visual acuity with no colour deficits.

*Apparatus.* The stimuli were presented on a 15-inch colour monitor with a refresh rate of 85 Hz. Trial events were controlled using MatLab software. Responses were made using the keyboard.

*Stimuli.* The fixation display consisted of a dot flanked by four squares (see Figure 1). At the viewing distance of 50 cm, the fixation point was 0.34 dva, and the place-holder squares were 1.15 dva. The eccentricity of the squares was 5.00 dva. All of these stimuli were grey against the black background.

When the cue was an onset, one of the squares in the cue display was briefly surrounded by four white dots (0.23 dva in diameter) in a diamond configuration. The distance from each dot to the edge of the square was 0.30 dva. When the cue was a colour, four purple dots surrounded one of the squares, while the other squares were surrounded by green dots.

The cue-and-target display consisted of the cue display with the addition of an “X” or an “=” (0.57 dva) inside one or all of the squares. For onset targets, only one square contained an “X” or an “=” (in white), while the other three were empty. For colour targets, two squares contained an “X” and two contained an “=”; only one of the characters was purple, while the others were green. Green and purple were selected to be equal in perceived luminance.

The target display was the same as above with the cues removed. This was followed by a return to the fixation display until a response was made.
Design. A 2 x 2 x 2 within-subject design was used. The factors were cue attribute, target attribute, and cue validity. Each combination of cue attribute and target attribute occurred in four contiguous blocks with the first block being practice. The order of conditions was counterbalanced. Cue validity varied randomly within block, and every combination of the four cue locations with the four target locations occurred equally often.

Procedure. The session lasted 1 hour, beginning with instructions that asked participants to be fast and accurate, and to maintain fixation throughout every trial. The uninformative nature of the cues was included in the instructions.

Each block began with two warm-up trials, followed by 32 data-collection trials, with two recovery trials inserted after an error. As shown in Figure 1, a trial began with the fixation display for 1000 ms, followed by the cue display for 35 ms, the cue-and-target display for 15 ms, and then the target display for 35 ms. The target display was followed by the fixation display until a response was made or 1500 ms had elapsed. The intertrial interval was 1000 ms. At the end of each block, participants were given summary feedback in terms of mean RT and accuracy.

All participants responded to “X” targets using the “Z” key and to “=” targets using the “?” key. Errors included making the wrong response, pressing an inactive key, not responding within 1500 ms, and responding in less than 50 ms. In every case, the appropriate error message was presented for 1500 ms. The data from warm-up and recovery trials, and all trials in practice blocks were excluded from analysis.

Results

Mean RTs as a function of cue attribute (onset vs. colour), target attribute (onset vs. colour), and cue validity (valid vs. invalid) are presented in Figure 2. A 2 x 2 x 2 repeated-measures ANOVA on the mean RTs confirmed the apparent three-way interaction, $F(1, 35) = 84.89, p < .001$. The initial analysis also showed main effects of target attribute, $F(1, 35) = 33.17, p < .001$, and cue validity, $F(1, 35) = 41.32, p < .001$, as well as all three of the two-way interactions, all $F(1, 35) > 24.20$, all $p < .001$. Consistent with previous work, as well as the figure, the three-way interaction was first explored by conducting separate two-way analyses, divided by target attribute. For the onset targets, the interaction between cue attribute and cue validity was significant, $F(1, 35) = 4.72, p = .037$. The same was true for colour targets, $F(1, 35) = 73.15, p < .001$, although the specific pattern was reversed. The second follow-up analysis reanalysed the data in terms of cue congruence and cue validity, which also produced a two-way interaction,
The simple main effect of spatial validity for congruent cues was large (59.65 ms) and significant, \(t(35) = 9.47, p < .001\), while the same effect for incongruent cues was null (-1.86 ms), \(t(35) = 0.39, p = .700\).

Overall error rates averaged 7.19%. There was no evidence for a speed-accuracy tradeoff: Across participants, error rates by condition were positively correlated with mean RT, mean \(r = .28, t(35) = 3.55, p < .001\).

**Discussion**

The present results are inconsistent with Rapid Disengagement and provide additional support for Contingent Capture. Rapid Disengagement explains previous failures to find spatial validity effects with incongruent cues by claiming that 150-ms delays between cues and targets are sufficient for both pure bottom-up capture and withdrawal of attention from that location. But the present experiment has also failed to find a spatial validity effect with incongruent cues at the very short SOA of 35 ms. At the same time, and

![Figure 2](image.png)

*Figure 2.* Mean response times and error rates as a function of target attribute, cue attribute, and cue validity.
going beyond previous work using the Cue Congruence Paradigm and short SOAs (e.g., Folk & Remington, 1998), the present experiment succeeded in producing a spatial validity effect with congruent cues, adding weight to the null finding with incongruent cues.

One might be tempted to argue that the incongruent cues in the present experiment did capture attention and then released it so rapidly that no spatial validity effect was observed, even at a very short SOA. However, a large body of work, using several methods, suggests that 35 ms is insufficient for both a shift of attention and then disengagement (e.g., Logan, 2005; Moore, Egeth, Beglan, & Luck, 1996; Theeuwes, Godijn, & Pratt, 2004). Furthermore, in a variant of the Cue Congruence Paradigm, Folk and Remington (2006) have shown that flanker-like distractors, appearing at the same time and location as the cue, only cause a response-compatibility effect when the cue is congruent. Thus, after modifying the paradigm in order to avoid some issues that arise when cues and targets appear simultaneously, this other work has found no evidence of capture by incongruent cues at an SOA of zero, leaving no time at all for Rapid Disengagement.

And, yet, the Cue Congruence Paradigm is only one of the methods that has been used to study attentional capture. A majority of the evidence supporting bottom-up models, including Rapid Disengagement, has been found using the Additional Singleton Paradigm (Theeuwes, 1992, 1994, 1996), instead. In a typical experiment using this method, an array of coloured shapes is presented, each surrounding a line; the task is to report the orientation of the line inside the one green diamond, ignoring the lines inside the (remaining) green circles (i.e., the defining attribute of the target is being inside a diamond). The critical finding concerns the effect of switching the colour of one nontarget circle from green to red. The slowing of RT on these trials, as compared with a baseline that does not include an additional singleton, is taken as evidence of bottom-up capture (since the colours of the items are task-irrelevant). Going further, Theeuwes et al. (2000) found evidence of capture using this task when the coloured shapes were presented at various SOAs before the oriented lines. In apparent conflict with the present work, distracting effects of an additional (incongruent) singleton were found at various SOAs, including 50 ms (see also Lamy & Egeth, 2003).

These data do not alter our conclusion in favour of Contingent Capture, as there are several ways to explain this discrepancy. First, in the additional singleton experiments with short delays between the coloured shapes and oriented lines, the SOAs were mixed within blocks, which has been shown to evoke the strategy of searching for any odd item (see Lamy, 2005). Conversely, the original version of the task only employed an SOA of zero (i.e., simultaneous onset of the coloured shapes and oriented lines), which would induce a filtering cost (Kahneman et al., 1983). More importantly, the two paradigms are very different in that one could omit the cues from the
Cue Congruence Paradigm and the task could still be performed (since the target, itself, contains the target-defining attribute); in other words, the cues used in the Cue Congruence Paradigm are truly task-irrelevant. In contrast, the coloured shapes that surround the oriented lines in the Additional Singleton Paradigm must be included, as they are the only items that indicate which line is to be reported; thus, the “cues” that are employed in the Additional Singleton Paradigm are not task-irrelevant. If one assumes that attention can be set to process a certain aspect of stimuli (e.g., the outer contour) without being set for any specific feature (e.g., diamond shape), then the data from the Additional Singleton Paradigm are quite consistent with those from the Cue Congruence Paradigm.

In summary, the evidence to date supports the idea that task-irrelevant items can only capture attention when they contain an attribute that matches current, top-down, control settings. The alternative idea that any salient item can capture attention, with attention being rapidly disengaged from incongruent items, is not supported.

REFERENCES


*Manuscript received December 2005
Manuscript accepted March 2007
First published online May 2007*