MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Top-down Guidance of Visual Attention by Specific Shapes"

verfasst von / submitted by

Marlene Forstinger, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Psychologie UG2002
Betreut von / Supervisor:	Univ.-Prof. Dr. Ulrich Ansorge

Danksagung

Ich bedanke mich bei Univ.-Prof. Dr. Ulrich Ansorge, der mich durch seine eigene Begeisterung für kognitionspsychologische Forschung inspiriert und auch in mir die Freude und das Interesse an diesen Themen immer wieder aufs Neue geweckt hat. Ich bedanke mich bei Markus, mit dem ich jeden Gedanken teilen kann, der mir gezeigt hat, was alles in mir steckt und der mein liebster Begleiter auf diesem Weg ist. Ich danke meinen Eltern, die mich schon mein ganzes Leben lang unterstützen und mir dieses Studium ermöglicht haben. Ich danke ihnen für alles, was sie mir für meinen Weg mitgegeben haben und dass sie mit mir buchstäblich durch dick und dünn gegangen sind. Ich danke meinem Opa, der mich so selbstlos unterstützt hat, immer das Beste für mich wollte und mir so viele Werte vermittelt hat, die mich für immer prägen. Ich danke Katrin und Helene, dass sie die besten Freundinnen sind, die ich mir wünschen kann und für ihr Sein, das mein Leben so bereichert. Zuletzt bin ich froh, dass ich mich nicht aufgegeben habe und den Mut gefasst habe, meinen eigenen Weg zu finden und zu gehen.

Marlene Forstinger, März 2020

Top-down Guidance of Attention by Specific Shapes

Marlene Forstinger

University of Vienna

Student no.: 01407966

Abstract

We investigated whether specific shapes (triangle pointing downwards, triangle pointing upwards, square, and diamond) can guide visual attention in a top-down manner using a contingent capture paradigm. Our results indicate that top-down attentional control settings for specific shapes can be established. Reaction times were significantly faster in trials when the target was presented at the same position as the preceding matching cue (valid trials) compared to when the matching cue preceded the target at a different position (invalid trials). In contrast, reaction times were not significantly different in valid and invalid trials with non-matching cues. This finding indicates that attention was only captured by cues that shared the target feature. The results of analyzing each target shape condition separately corresponded to our main analysis (which included the data of all four target shapes) except for the diamond target condition. While we found pronounced contingent capture effects in the other three target shape conditions, the cueing effect difference between matching and non-matching cues was scarcely evident in the diamond target condition. As it was already shown that similar stimuli capture attention as well, this result indicates that the diamond is similar to the hexagon (both shapes share four similar oblique lines). There was no speed-accuracy trade-off evident in the error rates.

Keywords: visual attention, shapes, top-down, contingent capture of visual attention, preattentive features

Contents

In every moment, we are simultaneously exposed to countless impressions in our environment. When we take a walk through the streets of a city, we can see people crossing streets and passing cars. We smell the scent of fresh bread from a bakery and hear birds chirping in the nearby trees. Regarding the visual system, the processing capacity is restricted in many ways. On the one hand, restrictions of the visual processing result from limitations arising of the physical properties of the eye as well as the visual system (e.g., visual accuracy is best in the fovea whereas vision is less accurate in the periphery of the retina).

On the other hand, it underlies other physical constraints, such as a limited wavelength range that can be processed (Carrasco, 2011). Thus, different selection mechanisms are necessary that enable us to distinguish relevant stimuli from irrelevant ones despite various limitations of the visual perception. This selectivity of our visual system is referred to as visual attention, which operates through different mechanisms that control attentional allocation (Awh, Belopolsky, & Theeuwes, 2012).

One of these mechanisms is bottom-up control, which describes involuntary, exogenous, and extrinsic shifts of attention that are elicited by the physical salience of stimuli and is also referred to as stimulus-driven control operating regardless of observers' intentional goals (Jonides, 1981). According to Itti, Koch, and Niebur (1998), salience of a stimulus is computed by three early visual features (color, orientation, and intensity) and visual attention is directed to the most salient location where the combined activity of these three features is highest. Perceptual salience of stimuli depends on the surrounding context. Thus, it may or may not be salient in different arrangements of objects (Itti & Koch, 2001). For example, Duncan and Humphreys (1989) showed that increasing target-distractor similarity impedes search while increasing distractor-distractor similarity facilitates search because then the target pops out. However, it would be hard to find a non-salient target if attentional allocation was solely bottom-up controlled by the salience of the stimulus. Thus, selection mechanisms that are controlled by current selection goals seem to be necessary to guide attention to non-salient targets. These mechanisms are often known as top-down or goal-driven, which refers to the initiation of endogenous, voluntary, and intrinsic attentional shifts (Wolfe, 1994). For example, in an early study, Helmholtz (1894) provided evidence for voluntarily controlled attention by showing that attention could be oriented without moving the eyes. According to Helmholtz (1894), this so-called covert attention leads to an enhanced perception of stimuli presented at the focused position, which was later also shown by Posner (1980).

However, attentional control is rarely exclusively determined by either bottom-up or top-down factors but is instead an interplay between the observer's intentions and the properties of the stimuli (Awh et al., 2012; Folk, Remington, & Johnston, 1992).

The mechanisms underlying attentional control were often investigated by using the spatial cueing paradigm (e.g., Posner & Snyder, 1975; Posner, 1980; Yantis & Jonides, 1990) in which the target is preceded by a cue that is presented either central of the display (e.g., an arrow pointing left) or peripheral (e.g., a red disk directly at or near the indicated position). The cue can appear at the same position as the following target or a different position as the target. While cues appearing at the same position as the target lead to facilitated processing of the target, which results in faster reaction times and more accurate responses, cues that appear at a different position as the target imped reaction times (Posner, 1980). Thus, measuring reaction times and error rates gives insight into the process of attentional allocation and whether cues capture attention. Jonides (1981) showed furthermore that peripheral cues influenced reaction times even though they were unpredictive of the position of the target, and participants were instructed to ignore the cue suggesting that attentional capture was involuntary, automatic, and bottom-up. The results by Yantis (1993) substantiated this finding by showing that peripheral visual information such as abrupt onsets is automatically selected regardless of the observers' intentions indicating that properties of the stimuli drive automatic attentional capture. However, the ability to suppress attentional shifts to stimuli suddenly appearing in our visual field is essential to prevent us from being distracted by irrelevant visual events. Yantis and Jonides (1990) showed that observers could override involuntary shifts of attention if observers are actively focusing their attention on a different location when the peripheral abrupt onset stimuli appear. Nevertheless, not only our active focus can modulate attentional allocation but also the observers' intentional goals.

According to the contingent capture hypothesis of Folk et al. (1992), attentional allocation is contingent on top-down attentional control settings that are induced by task demands. These top-down attentional control settings contain information about target-defining features and elicit involuntary and automatic attentional shifts to stimuli that share a critical feature with the target (e.g., the same color).

In the contingent capture paradigm, a modified spatial cueing paradigm, Folk et al. (1992) investigated attentional capture by using uninformative cues that either corresponded to participants' searched-for target feature (matching cues) or that did not share any critical feature with the target (non-matching cues). Attentional capture was measured by calculating a cueing effect meaning the difference between reaction times in trials with invalid targets (invalid trials) and trials with valid targets (valid trials). Folk et al. (1992) proposed that reaction times would be faster in valid trials compared to invalid trials with matching cues, whereas non-matching cues should not produce any cueing effect. As attentional capture by cues that appear at the same position as the following target facilitates target detection, reaction times are faster in valid trials compared to invalid trials in which target detection is impeded as attention has to reorient to a different than the cued location.

As predicted, Folk et al. (1992) found that uninformative cues automatically captured attention, but only if they corresponded to participants' searched-for target feature. For example, if participants searched for color targets, only color cues captured attention, whereas onset cues failed to capture attention. This finding also challenged the results by Yantis (1993) as onset cues did not produce automatic attentional shifts to their position if they did not correspond to the searched-for target feature.

Importantly, Folk et al. (1992) used cues that were unpredictive of the target's position, which means that only in 25 % of all trials, the cue appeared at the same position as the target (valid trials) whereas it preceded the target at a different position in 75% of all trials (invalid trials). As it was not useful for participants to shift their attention to the position of the uninformative cues, Folk et al. (1992) provided evidence for automatic and involuntary attentional capture contingent on top-down attentional control settings that were induced by participants' search goals. This interpretation of the results is strengthened because singleton color cues failed to capture attention, although they should have also caused involuntary attentional shifts to their position due to salience regardless of any search goals. Thus, the finding by Folk et al. (1992) shed new light on the conventional dichotomy between bottom-up and top-down controlled attentional capture by showing how observers' intentional goals interact with properties of the stimuli.

Initially, Folk et al. (1992) assumed that evidence for contingent capture effects could only be found with static features (e.g., color) and dynamic features (e.g., abrupt onsets). However, in later research, Folk and Remington (1998) found contingent capture effects furthermore for specific features (e.g., green) rather than a class of features (e.g., color) as long as a particular feature defines the target. While there is broad research about the influence of top-down attentional control settings for colors and onsets on attentional guidance (e.g., Ansorge, Kiss, Worschech, & Eimer, 2011; Folk et al., 1992; Goller & Ansorge, 2015), there is little research regarding shapes. However, in most natural search tasks, stimuli are not only defined by a unique color but also a specific shape. For example, if a man is searching for a rectangular emerald among differently shaped emeralds to surprise his wife for their anniversary, neither color information nor pop out will help him detecting the desired object as the rectangular emerald shares its color with the other emeralds and is not a singleton because it is presented among other shapes. Thus, the question arises whether shapes can control attentional allocation based on search goals. Therefore, in this study, we aim to extend the contingent capture hypothesis by investigating whether specific shapes can be used to guide visual attention top-down based on search goals.

Shapes as Guiding Attributes of Visual Attention

In the contingent capture paradigm, non-matching but salient cues do not capture attention, indicating that they are most probably discarded before attention is allocated to their position (but see, e.g., Theeuwes (2010) and Belopolsky, Schreij, and Theeuwes (2010) for alternative explanations of contingent capture effects). Thus, to find contingent capture effects for specific shapes, at least certain aspects of shapes must be processed before attention is allocated. Otherwise, cueing effects would not occur contingent on participant's searched-for target feature as attentional capture would be driven by the salience of the singleton cue. According to the original Feature Integration Theory by Treisman and Gelade (1980), preattentive processing was limited to a small set of basic features that are available in early vision (e.g., size, motion, orientation) and binding of basic features required attention. However, recent evidence suggests that certain aspects of shapes are also available in preattentive processing stages and, thus, can guide visual attention, but there is still uncertainty about what these aspects are (Wolfe & Utochkin, 2019).

In an early study, Wolfe and Bennett (1997) showed that an efficient search for a unique shape was not possible, concluding that global shape representation is not available in early vision. In contrast, Donnelly, Found, and Müller (2000) provided evidence for an efficient search for unique shapes if the target differed only by overall shape similarity from the homogenous distractors but not by component lines suggesting that some global shape information is available in early stages of visual attention. In line with Duncan and Humphreys (1989), Donnelly et al. (2000) proposed that the results by Wolfe and Bennett (1997) may have reflected impaired search by heterogeneous distractors instead of evidence against preattentive processing of shapes. However, although Donnelly et al. (2000) provided evidence for early global shape representation, the aspects of shapes that guide attention in visual search remain still elusive as there is recent evidence that the features available in early vision are not identical to those that are processed in a preattentive stage (Wolfe & Horowitz, 2017). Two specific criteria define the latter. First, they can guide visual attention, and second, they cannot be deconstructed into more basic features (Wolfe & Utochkin, 2019). Thus, although global shape information may be available in early vision, it may not be able to control visual attention as only if global shape information is preattentively processed it can guide attentional resources (Wolfe and Horowitz, 2017).

Regarding shapes, it is quite evident that they do not meet at least the second criteria as they are an albeit closed composition of some basic features (e.g., oriented lines) rather than an inseparable unity.

Thus, by investigating the contingency of attentional capture on top-down attentional control settings, we aim to provide further insight into the role of shapes in the guiding process of visual attention.

Current State of Knowledge: The Role of Shapes for Guiding Visual Attention

Up to now, the role of shapes for the guidance of visual attention was mainly investigated by studies investigating the dual-target theory (e.g., Adamo, Wozny, Pratt, & Ferber, 2010; Berggren & Eimer, 2018; Biderman, Biderman, Zivony, & Lamy, 2017) with shapes being the second target dimension next to color or in search for composite shapes (e.g., McCants, Berggren, & Eimer, 2018).

The results of Adamo et al. (2010) suggested that two top-down attentional control settings from different dimensions (shape and color) can be maintained in parallel. Although Adamo et al. (2010) indicated that shapes could be used to guide visual attention based on search goals, their results are limited due to their choice of stimuli. As they used only edgy and round shapes, it cannot be concluded that global shapes rather than information about edges guided attention, as the latter would have been sufficient to detect the target. To test this issue, we used simple polygons as targets that shared individual contours but were distinct in their global shape.

Further evidence for top-down attentional guidance by shapes was provided by Berggren and Eimer (2018), who showed that shapes guided attention top-down based on search goals, albeit not as efficient as colors. By using targets defined by a specific color and a specific shape, they found that distractors captured attention if only the shape matched the target, but the color was different, indicating that top-down attentional control settings were not only induced by search goals for specific colors but also for specific shapes.

Similar to Adamo et al. (2010) and Berggren and Eimer (2018), Biderman et al. (2017) were interested in simultaneous and selective attentional guidance by two features from different dimensions (shape and color), but their experiments also included single set conditions with the target being either a shape or a color. Interestingly, they did not find significant cueing contingent capture effects when the target feature was a shape.

Furthermore, McCants et al. (2018) showed that configuration search for shapes is initially part-based but determined by the global shape in later stages of attention. They found attentional capture not only for shapes that fully matched participants' searched-for target but also for partly matching shapes, indicating that closure might be a guiding feature of attention. Attentional capture for fully matching shapes was higher than for fully matching but reversed configuration shapes, suggesting that visual attention is furthermore sensitive to the arrangement of distinct global shapes that are defined by closed contours.

Altogether, there is contradicting evidence on whether shapes can guide visual attention based on top-down attentional control settings and on how efficient this guiding process is compared to other guiding attributes of visual attention such as color.

Therefore, we aimed to extend current knowledge about how shapes control attentional allocation. By using a contingent capture paradigm, we investigated whether attentional capture is contingent on search goals for specific shapes.

If task demands can induce top-down attentional control settings for specific shapes, we expect a significant cueing effect only for matching cues, whereas cues that do not share the target feature should not capture attention. If attentional allocation cannot be guided by top-down attentional control settings for specific shapes, we expect attentional capture for both the matching and the non-matching cue as they should elicit cueing effects of an equal or similar size driven by salience because the cue is always a singleton. As we do not assume differences in the guiding process of attention for different shapes, we expect similar results for each target shape. However, we will analyze the cueing effects for each target shape separately to confirm this assumption next to the analysis of all four target shapes together.

Method

Participants

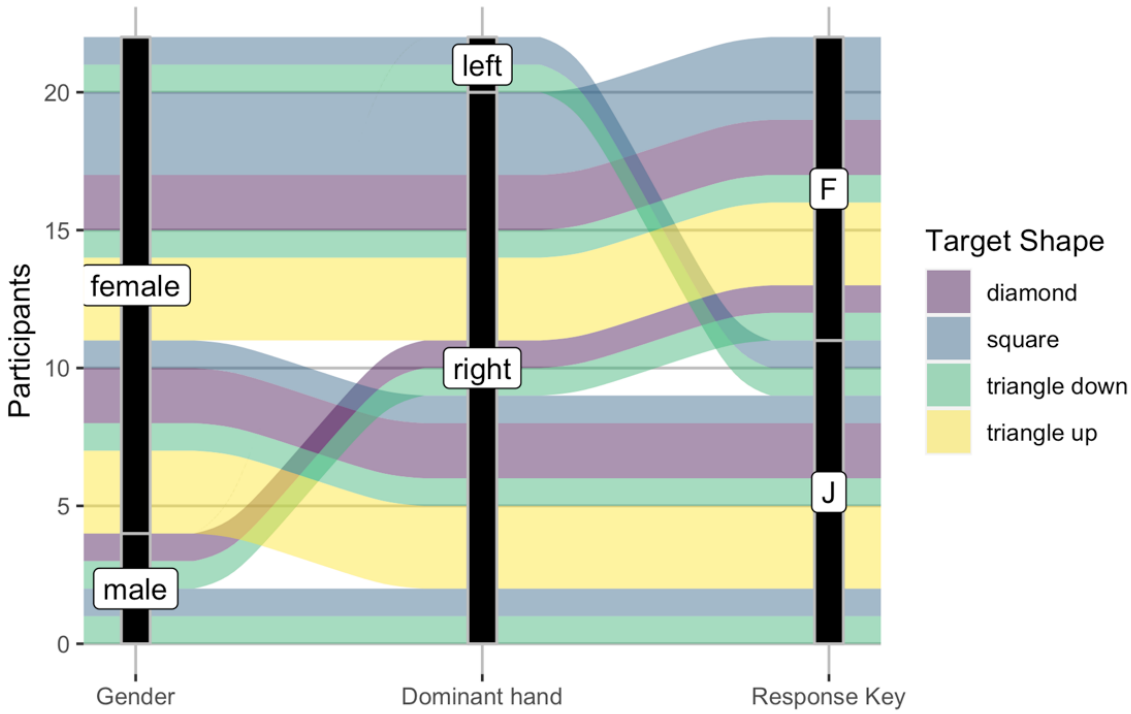
Twenty-five participants took part in this experiment. After excluding three subjects with more than 45% errors, 22 participants were analyzed (18 females; Mdn_{Age} = 19.5 years, range 19–24 years). All participants were naïve as to the purpose of the experiment and had normal or corrected to normal visual-acuity, no red-green deficiency, and gave written informed consent before the experiment. The experiment was conducted during a regular break in another, unrelated experiment for which participants received course credits. Further information about the sample is shown in Figure 1.

Apparatus and Stimuli

The experiment was conducted in a dimly lit room. Participants were instructed to use a chin-forehead-rest to maintain a stable viewing distance at 57 cm and to use a hearing protector to prevent distraction from possible background noise. The stimuli were generated and controlled with Matlab (MathWorks Inc., Natick, MA, USA) and the Psychophysics Toolbox (Brainard, 1997) and displayed on an LDC monitor. Stimuli displays (cueing display, masking display, and target display) consisted of four stimuli that were presented at the corners of an imaginary rectangle around the centroid of the screen with a resolution of 1.280 x 1.024 pixels and a refresh rate of 60 Hz. All stimuli were

Figure 1

Information About the Sample



Note. Information about the sample of the experiment (22 participants).

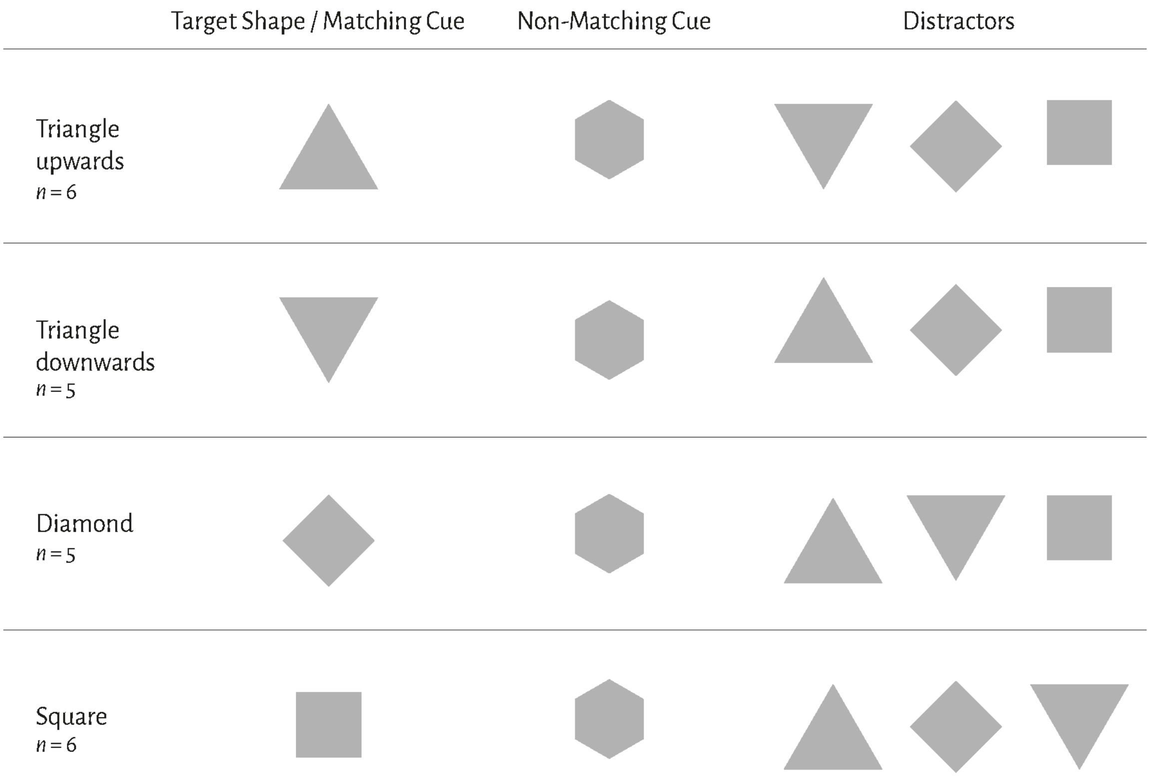
presented at the corners of an imaginary rectangle around the centroid of the screen with a resolution of 1.280 x 1.024 pixels and a refresh rate of 60 Hz. All stimuli were presented with a horizontal and vertical offset of 6.2° from the centroid and were displayed against a grey background (CIELAB color space: L* = 70, a* = 0, b* = 0). All stimuli were white (L* = 120, a* = 0, b* = 0). The area of all four stimuli in the target, as well as the cueing display, was equal. In the target display, all possible target shapes (triangle pointing upwards, triangle pointing downwards, square, and diamond) were presented simultaneously, but only one of them was the target shape. At the centroid of each target shape was a red (L* = 70, a* = 94, b* = 92) or green (L* = 70, a* = -79, b* = 51) disk with a diameter of 0.9° visual angle. The cueing display consisted of either one of the target shapes depending on the target shape condition (matching cue) or a hexagon (non-matching cue) and three white cue distractor disks. The side of the triangle pointing downwards and the side of the triangle pointing upwards was 4.33°. The side of the square and the side of the diamond was 2.85°. The hexagon had a side length of 1.77°. The cue distractor disks in the cueing display had a diameter of 3.22°. The size of the masking disks (diameter 5.3°) corresponded to the smallest size that was necessary to cover each of the four stimuli in the cueing display. The stimuli used in this experiment are shown in Figure 2.

As the shapes should be as simple as possible to keep the task difficulty low, we made four target shapes out of two. Thus, the triangle pointing upwards is a triangle pointing downwards that is 180° rotated, and the diamond is a square that is 45° rotated. However, as we investigated top-down guidance by specific shapes, the question arises whether the difference in orientation between two shapes influenced the search for the target shape.

According to Attneave (1968) and Humphreys and Quinlan (1988), global shape representations remain constant across transformations of orientation only if shapes have an unambiguous principal axis such as an isosceles triangle. If a shape has an ambiguous principal axis such as a square or an equilateral triangle as used in our study, changes in orientation elicit different perceptions (Palmer, 1985).

Figure 2

Illustration of the Stimuli Used in the Experiment



Note. Matching cues corresponded to the assigned target shape. The hexagon was a non-matching cue in each target shape condition. The distractors were shapes of the other target shape conditions, so that target displays of each target shape condition consisted of the same four shapes. In the experiment, the shapes were white ($L^* = 120$, $a^* = 0$, $b^* = 0$) and differed only in luminance from the grey background ($L^* = 70$, $a^* = 0$, $b^* = 0$).

Procedure

In this study, we used a contingent capture paradigm as we aimed to investigate whether the contingent capture hypothesis can be extended to the dimension of shape.

Additionally, the contingent capture paradigm provides advantages over the visual search paradigm, which has been commonly used in attentional research with shapes so far (e.g., Donnelly et al., 2000; Found & Müller, 1997; Wolfe & Bennett, 1997). For example, by using the contingent capture paradigm effects of confounding factors such as target and distractor similarity, which can negatively affect visual search efficiency

(Duncan & Humphreys, 1989) can be circumvented as they do not influence the cueing effect that is used to measure attentional capture.

As in visual search experiments, target and distractor are always presented simultaneously. Thus, one the one hand, target-related activity is enhanced to facilitate target detection, and on the other hand, distractor related activity is suppressed. As a consequence, suppression processes and facilitatory processes are in direct competition (Pinsk, Doniger, & Kastner, 2004). These processes can be decoupled in the contingent capture paradigm because the primary measure of attentional capture is the cueing effect, which is elicited by stimuli in the cueing display and not the time required for target detection. Additionally, suppression of distractors is held constant by using homogenous distractors, which reduces possible noise in attentional capture effects that would be stronger if heterogeneous distractors were used. Thus, by calculating the cueing effect, it can be investigated whether attentional capture is top-down guided by shared features between target and cue or properties of the stimuli (e.g., the salience of the cue). Additionally, as the cueing effect is calculated as the difference between reaction times in invalid and valid trials, the effect under investigation does not depend so much on the task difficulty or individual differences in reaction speed but instead on the reaction time difference between trials with cues that match or do not match a top-down attentional control setting.

We instructed participants to search for one of the four shapes (a triangle pointing downwards, rectangle, diamond, or triangle pointing upwards) that were simultaneously presented in the target display and report the color of the disk at its centroid (green or red). Participants used their right and left index fingers to press predefined keys on the keyboard (the J key for red and the F key for green or vice versa) to respond whether the disk in their target shape was red or green. Stimulus-response mappings were constant for each participant across the whole experiment but were counterbalanced between the participants. The target could appear at one out of four possible positions. The three target distractors also had a green or a red disk at their centroid and appeared at the three other positions.

Before data acquisition started, participants had to complete 20 practice trials. The real experiment only started if participants had reached more than 80% correct answers and consisted altogether of 576 trials with one self-paced break after 288 trials.

Each trial started with a fixation display (750 ms). Then the cueing display was presented for 50 ms, which was directly followed by the masking display (100 ms) and then the target display (350 ms). The target display was followed by a fixation display that was presented until participants responded via keypress or timeout was reached after 1 s. If participants reached timeout, they received the feedback “Zu langsam, bitte schneller reagieren” (too slow, please react faster). The next trial started after 500 ms intertrial interval again with a fixation display.

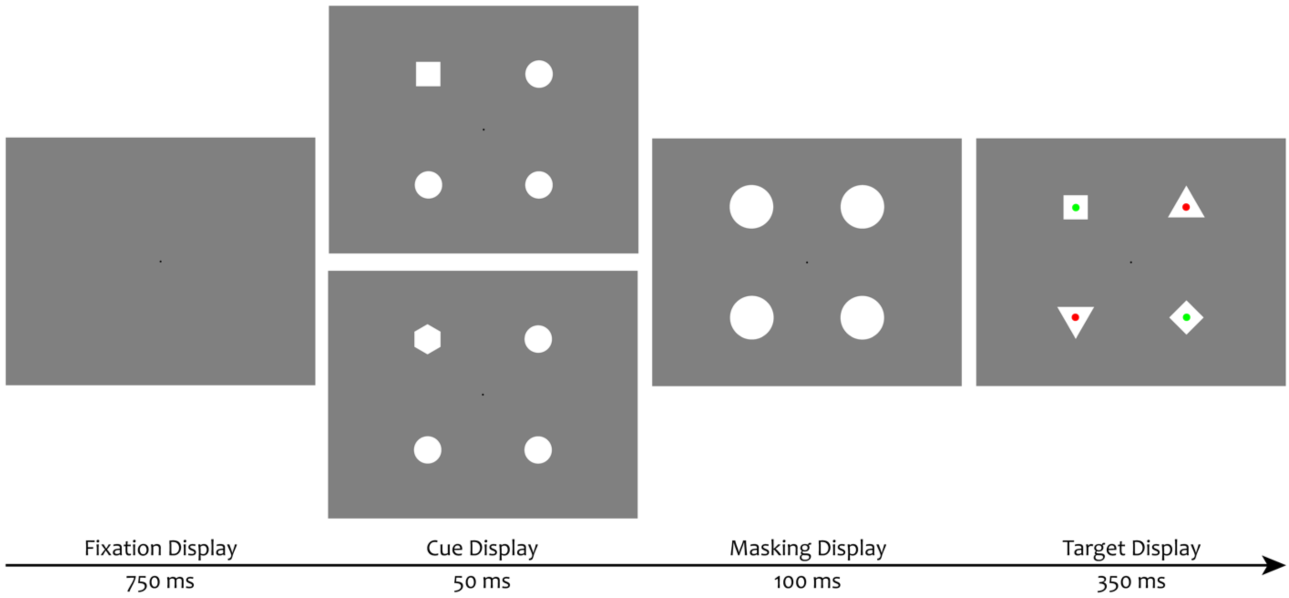
The cueing display consisted of four stimuli of which one was the cue, and the other three were always disks. In half of all trials, the cue matched participants’ target shape (a matching cue’s shape corresponded to the target shape), whereas, in the other half of all trials, the cue was a hexagon in all target shape conditions and thus did not match any target shape. The cue appeared at each of the four possible positions equally often. The cue and target positions were the same in 25% of all trials (valid trials), whereas their positions were different in 75% of all trials (invalid trials).

The masking display consisted of four white disks at each of the four stimulus positions of the previous cueing display. The masking screen was initiated to prevent perceptual facilitation by repeating the target position with a previous cue at the same position.

The target display consisted of four shapes (square, diamond, triangle pointing upwards, triangle pointing downwards). Only one of those four shapes was the target, which was counterbalanced between the participants and randomly assigned to the participants. The four shapes appeared equally often at each of the four possible target positions. In each centroid of the shapes in the target display was a colored disk. Two of the colored disks were always red, and two were always green. Each combination of red and green disks appeared equally often in the target display. An example of a valid trial of the square target condition is shown in Figure 3.

Figure 3

Procedure of the Experiment



Note. True to scale illustration of a valid trial of the square search condition. Both versions of the cueing display are depicted. The matching cue corresponded to the target shape and the non-matching cue was always a hexagon. Participants had to report the color of the disk in the middle of the target. In the illustrated trial, participants had to press the key for green to report the correct color of the disk in the square.

Data Analysis

We used R (Version 3.6.2; R Core Team, 2019) and the R-packages *data.table* (Version 1.12.8; Dowle & Srinivasan, 2019), *effsize* (Version 0.7.8; Torchiano, 2019), *emmeans* (Version 1.4.4; Lenth, 2019), *ggplot2* (Version 3.3.0.9000; Wickham, 2016), *lme4* (Version 1.1.21; Bates, Mächler, Bolker, & Walker, 2015), *lmerTest* (Version 3.1.1; Kuznetsova, Brockhoff, & Christensen, 2017), and *papaja* (Version 0.1.0.9942; Aust & Barth, 2018) for data analysis.

We used a 2 x 2 design with cue shape (matching vs. non-matching) and target validity (valid vs. invalid) as independent variables. Reaction times and error rates were dependent variables.

Results

We excluded reaction times below 150 ms and over 1 s (3.01%) from further data analysis and only analyzed reaction times of correct trials (15.5% wrong answers excluded).

Reaction Times

Reaction times were analyzed by using a linear mixed-effect model. We included cue shape and target validity as fixed factors, and random by-participants intercepts as a random factor. Additionally, we included interactions between the fixed factors.

Kliegl, Wei, Dambacher, Yan, and Zhou (2010) recommended using a linear mixed effect model for data analysis instead of a repeated-measures ANOVA for several reasons. Of particular interest for the contingent capture paradigm and the spatial cueing paradigm, in general, is that a linear mixed-effect model takes into account that there is a difference in the number of valid and invalid trials. According to Kliegl et al. (2010), a repeated-measures ANOVA ignores this imbalance in the number of trials between the two cue conditions that results in statistical power differences. Thus a linear mixed-effect model accounts for a more reliable measure, especially in the case of spatial cueing experiments and, therefore, also if the contingent capture paradigm is used.

Comparison of the main model including only fixed factors (cue shape and target validity) and the interaction model including an interaction between cue shape and target validity showed that the data were best described when the interaction was included. There was a significant interaction between cue shape and target validity, $b = 0.43$, 95% CI [0.03, 0.05], SE = 0.06, $t = 7.79$, $p < .001$ which was driven by faster reaction times in valid trials with matching cues (mean reaction time = 602 ms) compared to invalid trials with matching cues (mean reaction time = 644 ms), $t(21) = 6.33$, $p < .001$. For trials with non-matching cues, we did not find a significant difference between valid trials (mean reaction time = 631 ms) and invalid trials (mean reaction time = 629 ms), $t(21) = -0.25$, $p = .806$. The results of the linear mixed-effect model are shown in Table 1. The mean reaction times of correct trials for each experimental condition are shown in Figure 4.

Table 1

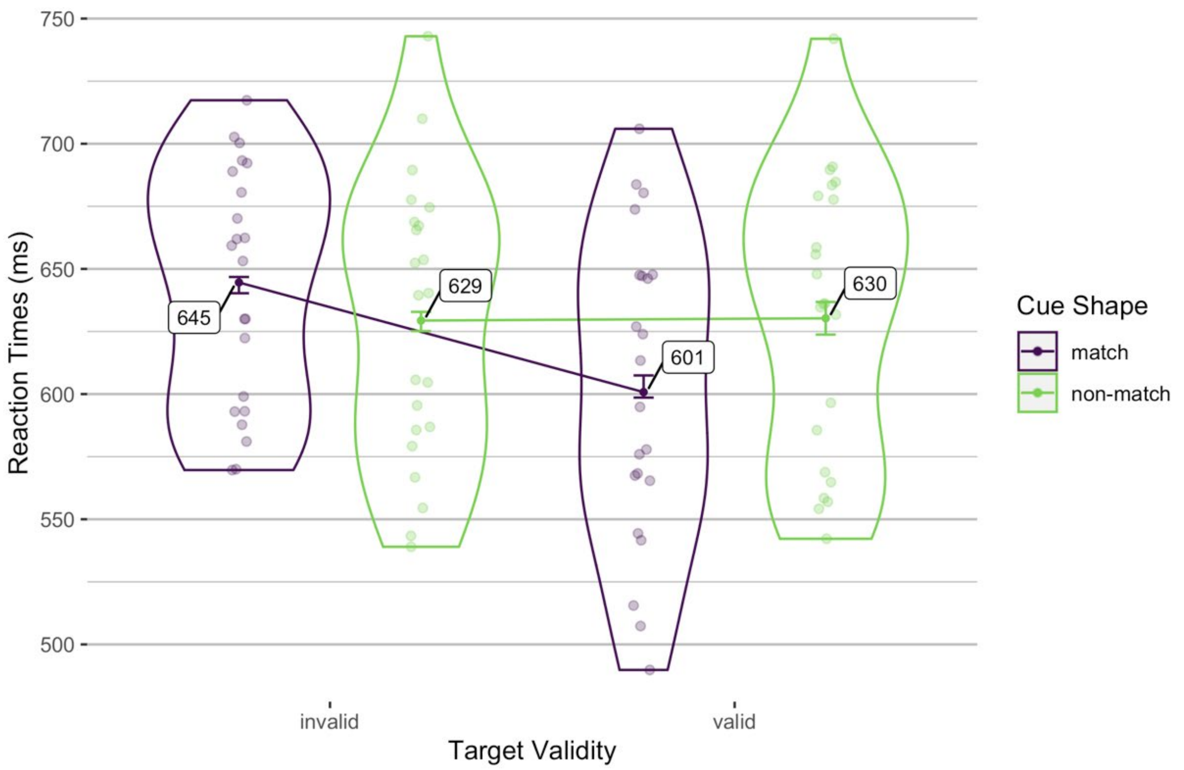
Results of the Linear Mixed-Effect Model

Fixed Effects	Estimate	95% CI	*SE*	*p*
intercept	643.80	[621, 666]	11.22	<.001
cue shape	-14.60	[-20, -9]	2.80	<.001
target validity	-42.53	[-50, -35]	3.93	<.001
cue shape * target validity	43.14	[32, 54]	5.54	<.001

Note. CI and *SE* in ms.

Figure 4

Mean Reaction Times of Correct Trials of Each Experimental Condition



Note. The semitransparent points represent individual mean reaction times, and the violin plot shows their distribution. The error bars represent 95% CI for the mean comparisons of all conditions (the 95% CI are adjusted by using the Tukey method for comparing a family of four estimates). The confidence interval is asymmetric for the largest and smallest value as it cannot be compared with a more extreme mean. The difference is significant if the error bars do not overlap.

Cueing Effect

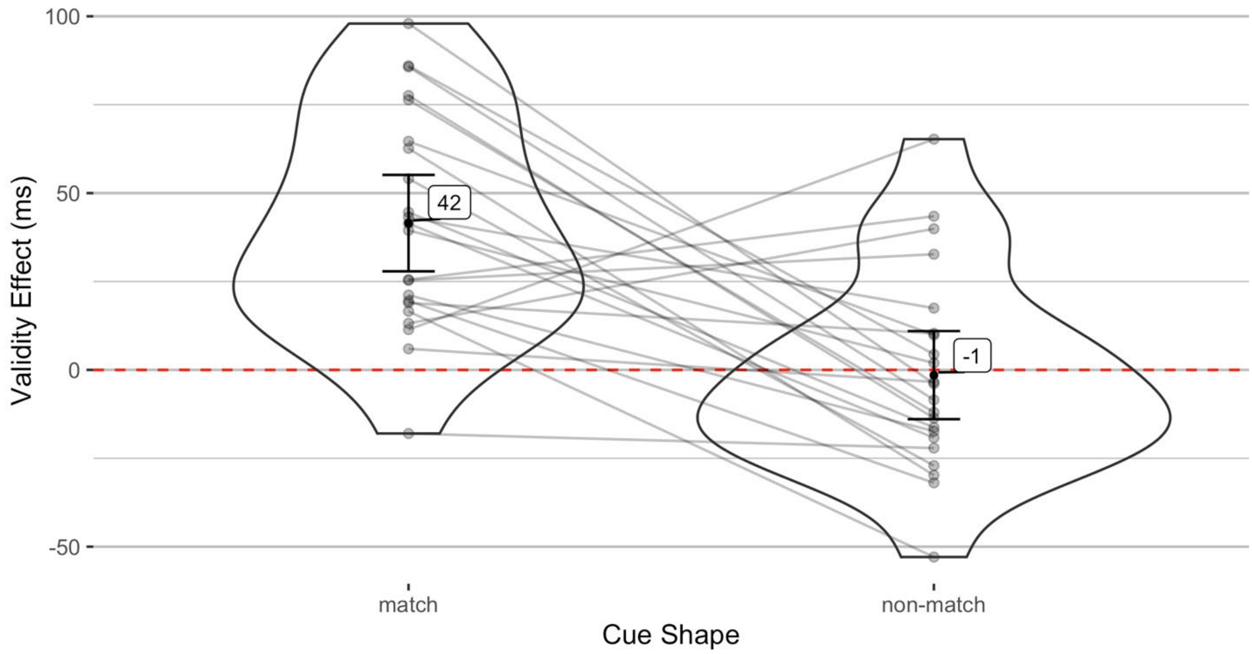
To test our hypothesis, whether attentional capture is contingent on search goals for specific shapes, we calculated a cueing effect by subtracting the mean reaction times of trials with invalid targets from trials with valid targets for trials with matching and non-matching cues separately. A positive cueing effect indicates attentional capture by a cue because of faster reaction times in valid trials due to facilitated processing of the target compared to slower reaction times in invalid trials as attention has to be reoriented to a position different from the cued position. We used a one-sample t-test to test whether matching and non-matching cues produced a cueing effect that was significantly different from zero (for all analyses, the significance level was set at $\alpha = 0.05$). A Shapiro-Wilk test for normality of the cueing effect's distribution in the matching cue condition showed that it was normally distributed, W = 0.96, $p = .507$. Also, for non-matching cues, the cueing effect's distribution was normally distributed, W = .96, $p = .497$. The cueing effect was significantly different from zero for trials with matching cues ($M = 42$ ms, $SD = 31$ ms, 95% CI [28, 55], $t(21) = 6.33$, $p < .001$, $d = 1.35$, 95% CI [0.76, 1.92]. In contrast to the matching cue, the cueing effect did not significantly differed from zero for trials with non-matching cues ($M = -1$ ms [-14, 11], $SD = 28$ ms, $t(21) = -0.25$, $p = .806$, $d = -0.05$, 95% CI [-0.47, 0.37]). Additionally, we tested whether the cueing effect was significantly different in trials with matching cues compared to non-matching cues. A Shapiro-Wilk test showed that the cueing effect difference was normally distributed (W = 0.94, $p = 0.186$) and a two sided t-test for paired samples showed that the cueing effect difference was significant ($M = 43$ ms [23, 63], $SD = 45$ ms, $t(21) = 4.51$, $p < .001$, $g = 1.90$, [0.73 1.90]. This effect was driven by a stronger cueing effect in trials with matching cues (42 ms) compared to trials with non-matching cues (-1 ms). Figure 5 shows the mean cueing effect difference for matching and non-matching cues from zero ms with error bars representing 95% CI.

Error Rates

For analyzing the error rates, we used a generalized linear mixed-effect model with a binomial link function to test whether the reaction times in valid trials with matching

Figure 5

Cueing Effect for Each Cue Condition

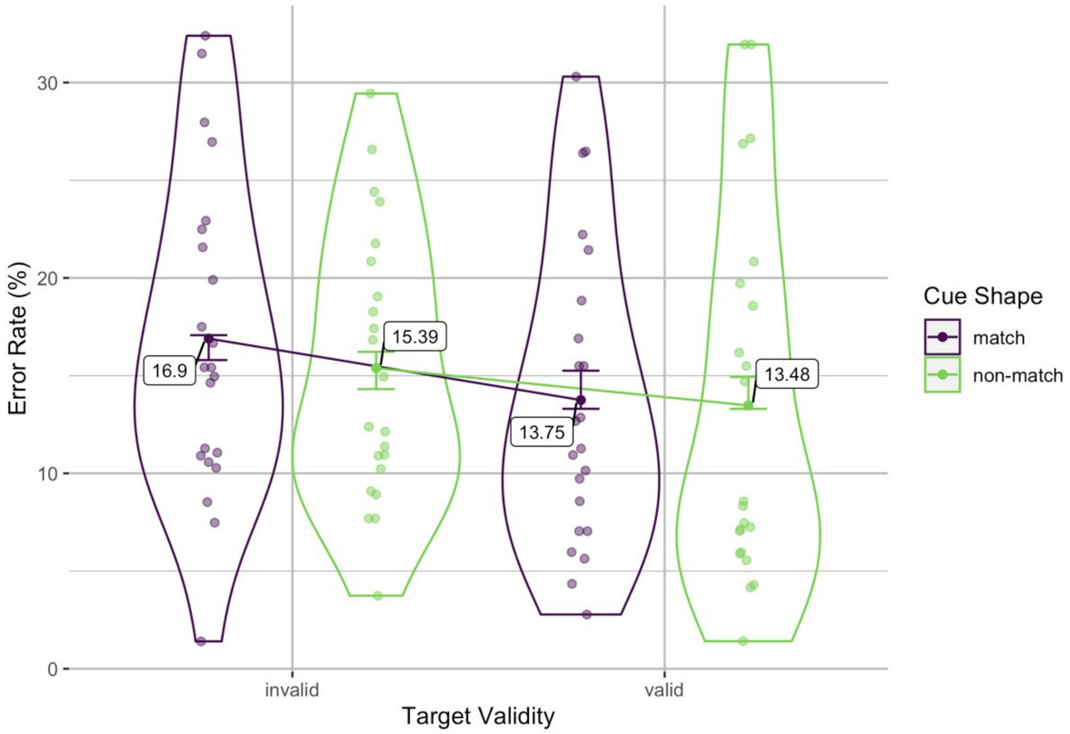


Note. The semitransparent points represent the individual validity effects, and the violin plot represents their distribution. Lines connect paired values, and the error bars represent the 95% CI for the one-sample t-test against zero (red dashed line). The difference is significant if the error bars do not overlap.

cues were only faster at the expanse of accuracy. Hierarchical model comparisons showed that the main model, including target validity as a fixed factor and random-by-participants intercept as a random factor, fit the data best. We found a significant error rate difference between valid and invalid trials (2.41%), $z = 3.371$, $p < .001$. This effect was driven by a lower mean error rate in valid trials (12,31%) compared to invalid trials (14.72%). The mean error rates for each experimental condition are shown in Figure 6. Table 2 shows differences in mean error rates for each experimental condition.

Figure 6

Mean Error Rate for Each Experimental Condition



Note. The semitransparent points represent individual mean error rates, and the violin plot shows their distribution. The error bars represent the 95% CI for all mean comparisons. The CIs were adjusted by using the Tukey method for comparing a family of four estimates.

Table 2

Mean Error Rates

Cue Shape	Target Validity	Mean Error Rate (%)	95% CI	*SE*
match	invalid	16.9	[15.84, 17.96]	0.54
non-match	invalid	15.39	[14.37, 16.41]	0.52
match	valid	13.75	[12.07, 15.43]	0.86
non-match	valid	13.48	[11.83, 15.12]	0.84

Note. CI and *SE* in %.

Additional Analysis

For exploratory interest, we analyzed the mean reaction times for each target shape (triangle pointing downwards, triangle pointing upwards, square and diamond) separately in order to reveal possible cueing effect differences between the shapes. However, it must be considered that the mean reaction times for each target shape are based on only five to six participants. Thus, the results of this analysis must be taken with a grain of salt as the statistical power is questionable due to the small sample size.

First, we wanted to analyze whether there were general differences in reaction times between the shapes. Therefore, reaction times were analyzed by using a linear mixed-effect model, including target shape as a fixed factor and random by-participants intercepts as a random factor. Table 3 shows the mean reaction time for each target shape.

Furthermore, we were interested in whether there were general differences between the experimental conditions of the four target shapes. The mean reaction times of correct trials for each experimental condition, and each target shape are shown in Figure 7.

Additionally, we investigated differences in mean reaction times between the target shapes, which could indicate that search difficulty varied between the shapes. Comparing two target shapes at a time revealed no significant difference in the mean reaction times between any of the four target shapes. Table 4 shows the differences in the mean reaction times between two target shapes at a time.

Table 3

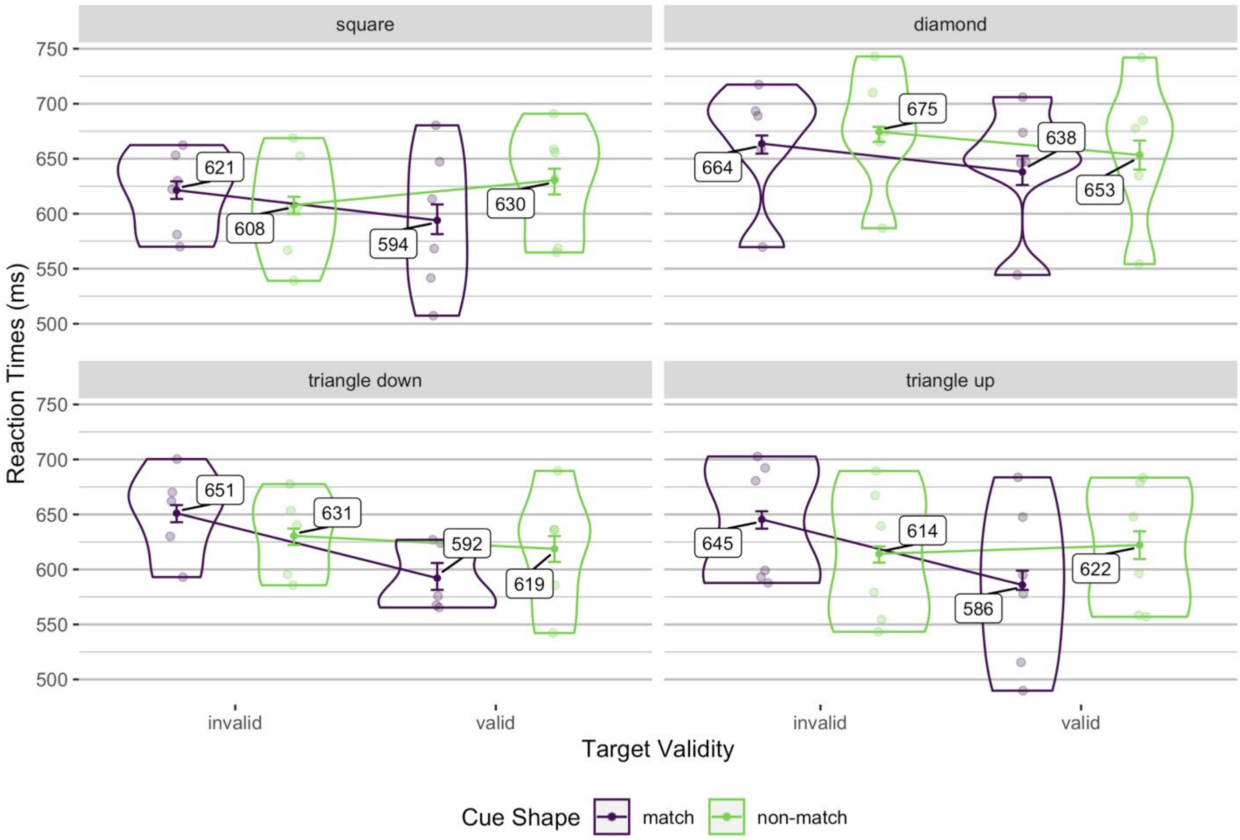
Mean Reaction Time for Each Target Shape

Target Shape	Mean Reaction Time (ms)	95% CI	*SE*
square	612	[571, 654]	21
diamond	665	[620, 711]	23
triangle down	632	[586, 677]	23
triangle up	621	[579, 662]	21

Note. CI and *SE* in ms.

Figure 7

Mean Reaction Times of Correct Trials for Each Experimental Condition, and Each Target Shape



Note. The semitransparent points represent individual mean reaction times, and the violin plot represents their distribution. The error bars represent 95% CI for all mean comparisons (adjusted using the Tukey method for comparing a family of four estimates). The confidence interval is asymmetric for the largest and smallest value since it cannot be compared with a more extreme mean. If the error bars do not overlap, the difference is significant.

Furthermore, we analyzed the mean error rates for each target shape separately to reveal possible differences in the mean error rates between the four target shapes. If search difficulty varied between the shapes, participants would have probably produced more wrong answers dependent on their assigned target condition. The mean error rates for each experimental condition and each target shape is shown in Figure 8.

Table 4

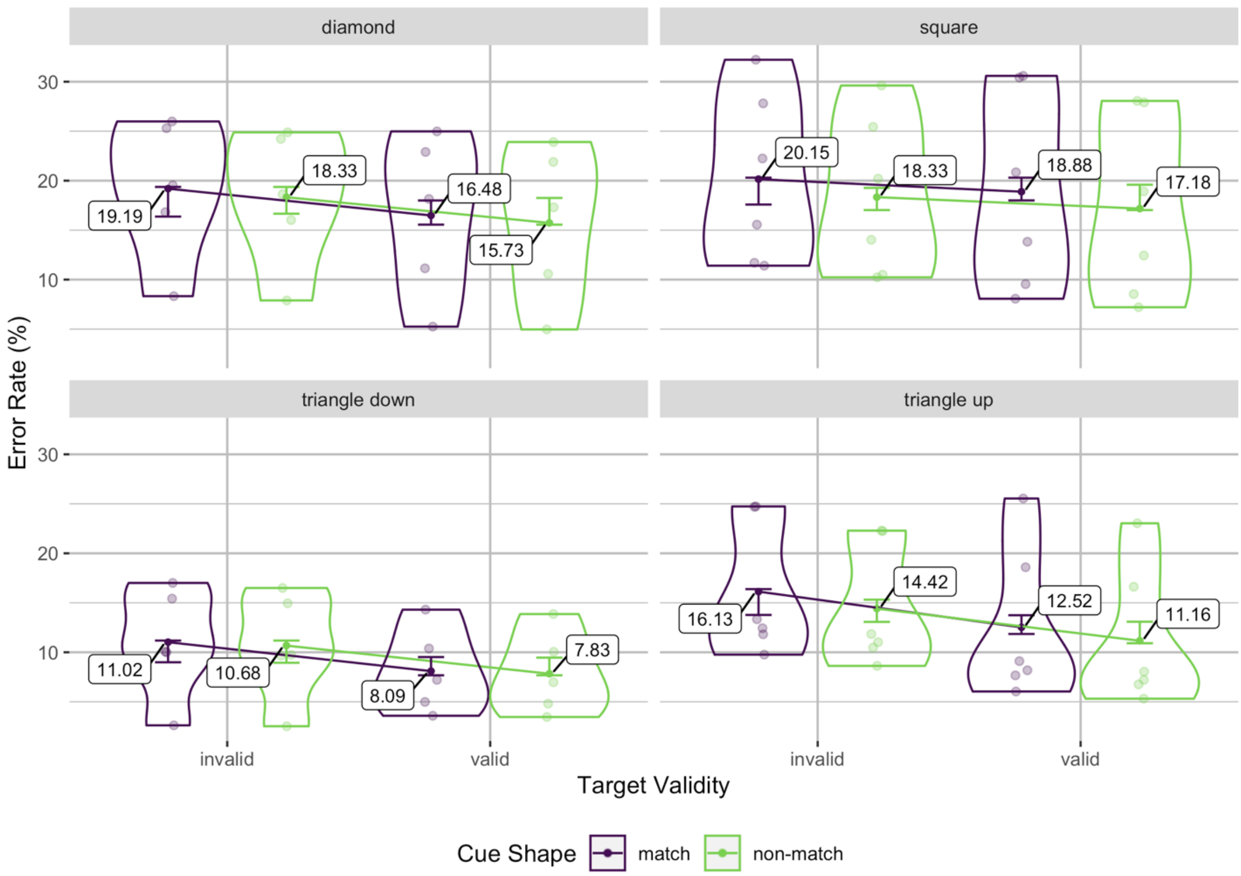
Mean Reaction Time Difference Between Target Shapes

Contrasts	Mean Difference (ms)	*SE*	z-ratio	*p*
square - diamond	-53.22	31.46	-1.69	.330
square - triangle down	-19.34	31.45	-0.61	.930
square - triangle up	-8.59	29.99	-0.29	.990
diamond - triangle down	33.88	32.85	1.03	.730
diamond - triangle up	44.63	31.45	1.42	.490
triangle down - triangle up	10.75	31.44	0.34	.990

Note. SE in ms.

Figure 8

Mean Error Rate for Each Experimental Condition, and Each Target Shape



Note. The semitransparent points represent individual mean error rates, and the violin plot shows their distribution. The error bars represent the 95% CI for all mean comparisons. The CIs were adjusted by using the Tukey method for comparing a family of four estimates.

To test whether the error rate depended on the target shape condition, we used a generalized mixed-effect model with a binomial link function, including target shape as a fixed factor and random-by-participants intercept as a random factor to investigate whether the mean error rates were different between the target shapes and found no significant difference. Table 5 shows the mean error rate for each target shape.

Additionally, we compared the mean error rates of two target shapes at a time to shed further light on possible differences in task difficulty between the shapes which may be indicated by differences in the mean error rates. Table 6 shows the differences in mean error rates between two target shapes at a time.

Table 5

Mean Error Rate for Each Target Shape

Target Shape	Mean Error Rate (%)	95% CI	*SE*
square	17.69	[11.74, 23.65]	3.04
diamond	17.16	[10.77, 23.54]	3.26
triangle down	9.13	[5.33, 12.94]	1.94
triangle up	13.5	[8.68, 18.32]	2.46

Note. CI and *SE* in %.

Table 6

Mean Error Rate Difference Between Target Shapes

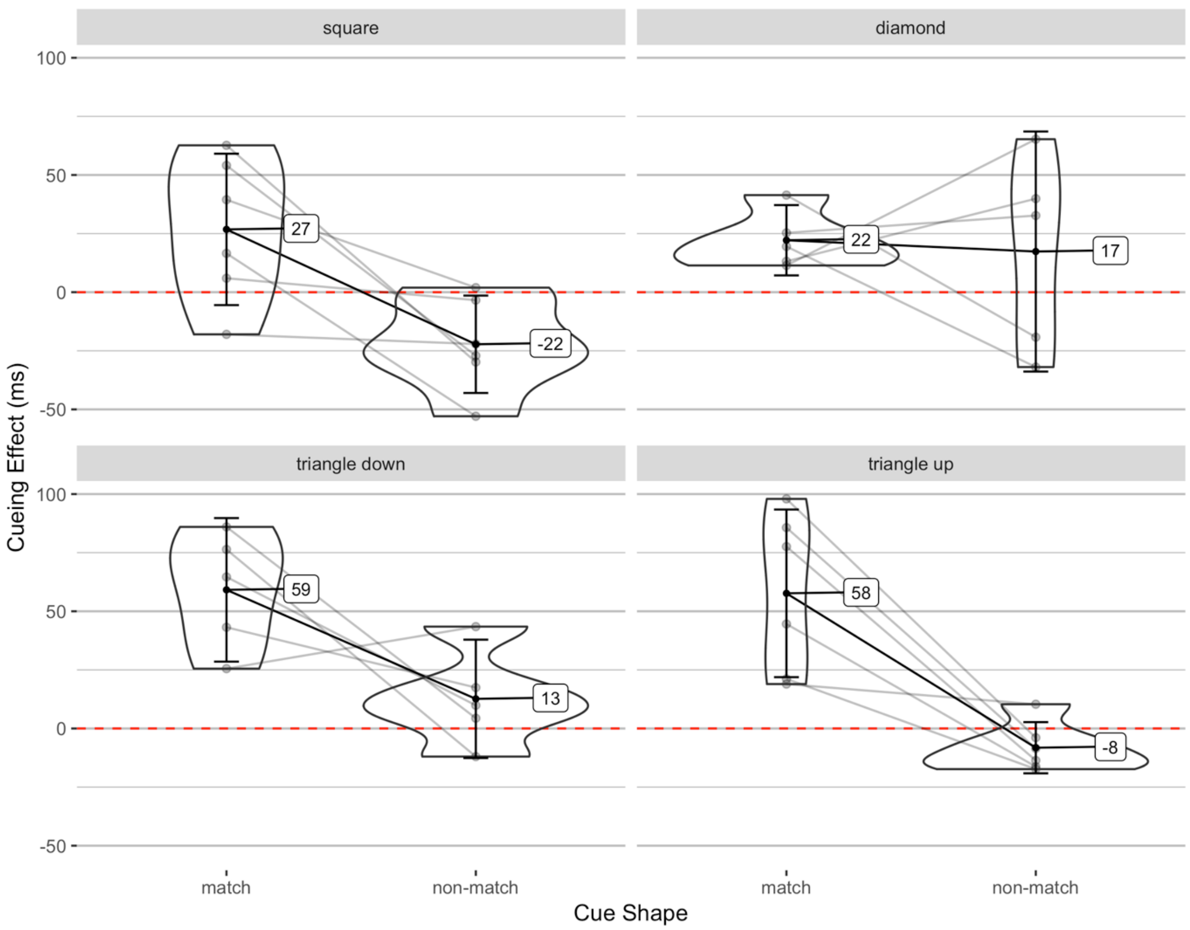
Contrasts	Mean Error Rate Difference (%)	*SE*	z.ratio	*p*
square - diamond	0.54	4.45	0.12	.999
square - triangle down	8.56	3.61	2.37	.082
square - triangle up	4.19	3.91	1.07	.706
diamond - triangle down	8.02	3.79	2.11	.148
diamond - triangle up	3.66	4.08	0.9	.807
triangle down - triangle up	-4.37	3.13	-1.39	.503

Note. SE in %.

Furthermore, we calculated a cueing effect for each target shape separately for matching and non-matching cues to find out whether the top-down attentional control settings for the different target shapes were equally efficient in guiding visual attention, which would be evident in cueing effect differences. The cueing effect for each cue condition and each target shape is shown in Figure 9.

Figure 9

Mean Cueing Effect for Each Cue Condition, and Each Target Shape



Note. The semitransparent points represent the individual validity effects, and the violin plot represents their distribution. The lines connect paired values, and the error bars represent the 95% CI for the one-sample t-test against zero (red dashed line). No overlap between the error bars indicates a significant difference. The difference between trials with matching and non-matching cues are as follows:
square (n = 6): 49 ms, t(5) = 3.20, p = .024, g = 1.32, 95% CI [0.02, 2.28],
diamond (n = 5): 5 ms, t(4) = 0.22, p = .839, g = 0.20, 95% CI [-0.84, 1.13],
triangle up (n = 6): 66 ms, t(5) = 4.38, p = .007, g = 1.86, 95% CI [0.27, 2.89],
triangle down (n = 5): 46 ms, t(4) = 2.38, p = .076, g = 1.44, 95% CI [-0.20, 2.19].

Discussion

In this study, we showed that simple shapes guide visual attention top-down based on search goals. As predicted by the contingent capture hypothesis, involuntary attentional capture of cues was contingent on task-relevant features as indicated by a significant cueing effect only for cues that shared the critical target feature (cues that had the same shape as the predefined target). If attentional capture had not been contingent on search goals, the non-matching cue (a hexagon in all target shape conditions) would have also produced a cueing effect driven by the cue's salience as a singleton in the cueing display. Analysis of the error rates showed no significant difference between trials with matching and non-matching cues, indicating that reaction times in valid trials with matching cues were not faster at the expense of accuracy.

Additionally, we analyzed each target shape separately. Taking a closer look at each target shape should reveal possible cueing effect differences between the shapes that could give insight into whether the efficiency of the top-down attentional control settings was similar for the four different shapes. Considering that the results of each target shape analysis are based on only five to six participants (hence the statistical power was probably low), they must be taken with a grain of salt.

Comparing the mean reaction times and the mean error rates between the four target shapes, we did not find a significant difference. Thus, our results indicate that the task difficulty was independent of the target shape condition as otherwise, participants would have probably produced more errors and longer reaction times in certain target shape conditions if the search for a specific shape was more difficult than for others.

Analyzing the cueing effect for each target shape separately, we found more substantial cueing effects for matching cues compared to non-matching cues in all target shape conditions except for the diamond target condition. Interestingly, only in the diamond target condition, the cueing effects in the two cue conditions were nearly equal in strength (matching cue = 22 ms, non-matching cue = 17 ms). Thus, the cueing effect difference was negligible (5 ms), whereas, in the other target shape conditions, the cueing effect difference was more pronounced. This finding in the diamond search condition suggests that not only the top-down matching cue (diamond) captured attention but also the top-down non-matching cue (hexagon). Although this effect may not be reliable as indicated by large intervals of confidence of the mean cueing effects for matching and non-matching cues, it may reflect activation of the top-down attentional control setting for the diamond not exclusively by the target shape itself but also by the hexagon. This effect may be driven by the similarity between these two shapes as the hexagon shares four similar oblique contours with a diamond. The hexagon resembles no other target shape to this extent, which could explain why this effect solely occurred in the diamond search condition. Alternatively, the diamond target condition could reflect bottom-up capture effects driven by the salience of the (matching and non-matching) cue in the cueing display. However, considering the unambiguous contingent capture effects in all other target shape conditions, this approach seems unlikely.

Interestingly, although a square shares two parallel contours with a hexagon, no similar cueing effect for the non-matching cue was found in the square target condition. If similarity explains the attentional capture by the non-matching cue in the diamond search condition, similar albeit smaller attentional capture effects in search of squares would be expected. Nevertheless, given our findings, similarity does not seem to be calculated solely as a function of shared contours but rather reflect that certain aspects of shapes contribute more to perceived similarity than others. As the results of analyzing each target shape are limited due to the small sample size, replication of this surprising finding in the diamond target condition with a higher statistical power is necessary to test whether this finding is reliable. However, as we found pronounced contingent capture effects in all other target shape conditions, our results indicate that different shapes are similar in their efficiency of guiding visual attention top-down based on search goals. Nevertheless, our results indicate that processing differences between shapes exist, and future research may investigate the aspects of shapes that underlie these discrepancies in the efficiency of controlling attentional allocation top-down.

The Disengagement Theory of Attention

The contingent capture hypothesis has been challenged by the disengagement theory of attention, which states that attention is initially captured not only by the matching cue but also by the non-matching cue due to the cue's salience in the cueing display. However, attention is rapidly detached from the task-irrelevant stimulus as soon as it is identified as not corresponding to any top-down attentional control setting, while in contrast, attentional disengagement from a matching cue is relatively slow (Theeuwes, 2010; Theeuwes, Atchley, & Kramer, 2000). Therefore, attentional capture can be measured solely for matching cues, which results in standard contingent capture effects. According to this account, top-down factors influence attentional allocation only in later processing stages. However, this was already contradicted by Ansorge et al. (2011), who showed that visual selection is controlled top-down from the very start of visual processing. Although numerous studies provided evidence against the disengagement theory (e.g., Ansorge et al., 2011; Chen & Mordkoff, 2007; Eimer & Kiss, 2008, 2010) it can hardly be falsified because the process of disengagement is said to be very fast and as a result, attentional capture by irrelevant stimuli is not evident in behavioral data (e.g., reaction times). Therefore, one commonly used way to test whether attention was rapidly removed from a non-matching cues' location after it initially captured attention due to its salience is to reduce the cue-target stimulus-onset asynchronies (SOAs). Using this method strictly encloses the period in which rapid disengagement can take place and, thus, gives insight into whether also the non-matching cue captures attention initially. The results of a meta-study by Büsel, Voracek, and Ansorge (2018) showed that rapid disengagement was not evident by using very short SOAs, indicating that rapid disengagement cannot account for contingent capture effects. Furthermore, neurophysiological data (which has a higher temporal resolution than behavioral measurements) was provided against the disengagement theory (e.g., Eimer and Kiss (2008); Eimer and Kiss (2010)). However, Belopolsky et al. (2010) proposed that neurophysiological studies cannot provide convincing evidence against the disengagement theory of attention as neurophysiological measurements may not reveal attentional capture but rather reflect the processing of the target while stimuli that are minimally processed (e.g., non-matching cues) will not be evident in the data.

Intertrial Priming Effects on Attentional Capture

Furthermore, alternative accounts of attentional capture propose that intertrial priming can explain or at least contributes to contingent capture effects (Belopolsky et al., 2010; Theeuwes, 2013). According to intertrial priming approaches of attentional capture, a matching cue captures attention because the feature that a matching cue shares with the target is primed by the previous target rather than because of activation of a top-down attentional control setting. As the target feature is constant across all trials in the standard contingent capture paradigm that we also used in our study, such intertrial priming effects cannot be ruled out. Whether the contingency of cueing effects can be explained by intertrial priming effects rather than by activation of top-down attentional control settings, has been investigated by using a mixed-target design in which the target feature changes unpredictably across trials (e.g., Ansorge & Horstmann, 2007; Folk & Remington, 2008). According to intertrial priming accounts, attentional capture of a cue is determined by its feature congruence with a previous target (e.g., Awh et al., 2012; Theeuwes, 2013). Thus, as the target feature changes unpredictably across trials in mixed-target designs, contingent capture effects should be negatively affected if intertrial priming determines attentional capture as the cue would only capture attention if the previous target primed its feature.

In contrast, if top-down attentional control settings for specific target features guide attentional allocation, cueing effects should be contingent on the feature congruence between cue and target instead of previous target features. By using a mixed-target design, Folk and Remington (2008) found that attentional capture was contingent on the feature match between a cue and the previous target only if top-down attentional control settings were unspecific thus, not predefined by instruction. Ansorge and Horstmann (2007) strengthened this finding by showing that cueing effects were contingent on top-down attentional control settings if they were specified in advance of the experiment. Additionally, Büsel et al. (2018) and Lamy and Kristjansson (2013) reviewed the evidence on studies investigating the intertrial priming contingencies of attentional capture and found converging evidence for top-down attentional control settings as the driving force of contingent capture effects. However, the influence of intertrial priming on attentional capture is diverse. For example, Goller and Ansorge (2015) tested a different possible influence of intertrial priming by investigating whether attentional capture of an unpredictive cue is affected by the usefulness of the previous cue, thus by whether a trial is valid or invalid. As suggested by Jongen and Smulders (2007), attentional capture may be influenced by the cue-target-relations of the previous trial suggesting that participants are encouraged to use an uninformative cue if it was useful in a previous trial. Interestingly, Goller and Ansorge (2015) found cueing effects for non-matching cues if the cue of the previous trials was valid, whereas attentional capture by matching cues was not contingent on previous cue-target-relations. According to Goller and Ansorge (2015), their results could suggest that expectancies regarding the non-matching cue's usefulness may have caused attentional capture contingent on the target validity of the previous trial, which was previously shown by Gratton, Coles, and Donchin (1992). However, the specific processes underlying these intertrial priming effects on attentional capture remain elusive and are subject to future research. Taken together, even though intertrial priming may contribute to contingent capture effects, there is converging evidence showing that it cannot fully explain the contingencies of attentional capture.

Future Research Implications

Top-down Guidance of Visual Attention by Global Shapes Versus Local Features

Our results correspond to our initial hypothesis that specific shapes can guide visual attention top-down, but alternatively, the effects could have also been driven by top-down attentional control settings for only basic features such as local contours, or angles instead of global shapes. Regarding the contour lines, it has to be noted, that the stimuli in our study solely differed in luminance from the grey background and thus, were not composed of contours lines. That way, we aimed to highlight the global shape rather than distinct local features as it has already been shown that oriented lines can guide visual attention (for a review, see Wolfe & Utochkin, 2019). However, the results by Found and Müller (1997) suggested that search for targets defined by cognitive contours, which mean boundaries emerging from differences in texture, was equally efficient than for targets defined by contour lines. This finding suggests that boundaries emerging from a difference in luminance, as evident in our study, were probably also perceived as cognitive contours and, therefore, could have been used to guide visual attention. However, as mentioned before, Donnelly et al. (2000) provided evidence for the representation of global shapes in early vision and their ability to guide attention efficiently in visual search. Thus, they could refute the earlier finding that shapes were represented as lose bundles of contour lines before attentional processing is engaged, as suggested by Wolfe and Bennett (1997). However, if local contours guided visual attention independently from the global shape they are part of, we would have probably also found pronounced cueing effects using only distinct contour lines such as two parallel vertical lines in the square condition as a matching cue and a hexagon as a non-matching cue. As two parallel vertical lines are solely evident in the square, they should capture attention if top-down attentional control settings can be established for target-defining local contours instead of or parallel to top-down attentional control settings for global shapes. By using this procedure with the matching cues adjusted correspondingly to each target shape, top-down attentional control settings for local contours versus global shapes could be investigated for other shapes as well. However, Donnelly et al. (2000) provided evidence against local contours as independent guiding features of attention in the search for shape targets and showed that early vision processing is sensitive to global shapes if their contours exhibit closure. Most importantly, in their third experiment, Donnelly et al. (2000) found that this early representation of a global shape cannot be segmented into its various simpler components by showing that global shape similarity between the stimuli in the target display was the critical factor explaining search efficiency irrespective of whether the target shape contained a unique contour or not. This finding strengthens the assumption that global shapes are an autonomous guiding feature of visual attention rather than solely a composition of local contours that are independently processed even in early stages of visual processing.

Sensitivity of the Visual System for Specific Orientations of Shapes

If retinal coded features determined early shape representations (e.g., contour lines, edges, and angles), they would be strongly affected by changes in orientation and size. However, if local features are coded relative to a reference frame such as an axis of symmetry, a consistent representation across changes in retinal projections (e.g., by changes in orientation or size) can be attained (Duncan & Humphreys, 1989; Humphreys, 1983; Humphreys & Quinlan, 1988). However, an axis of symmetry can only account for constant shape representation across different transformations if shapes have an unambiguous axis of symmetry, such as an isosceles triangle (Humphreys & Quinlan, 1988). Humphreys and Quinlan (1988) found that orientational transformations of shapes elicited different visual percepts if the shapes had an ambiguous axis of symmetry. They showed that reaction times were impeded if a target with an ambiguous axis of symmetry (e.g., a square) was primed by a shape that shared the target's global shape but differed in its orientation (e.g., a square rotated by 45°). In contrast, priming effects for shapes with an unambiguous axis of symmetry (e.g., a pentagon or an isosceles triangle) were unaffected by orientational differences between prime and target. The results by Humphreys and Quinlan (1988) suggested that such ambiguous shapes are only consistently represented across orientational transformations if the reference frame changes, respectively. Furthermore, regarding the contingent capture paradigm, this finding indicates that attentional capture may not only be contingent on whether cue and target have the same shape but may also be differently affected by whether cue and target are equally oriented depending on the ambiguity of their axes of symmetry. Therefore, future research may address whether top-down attentional control settings for shapes are sensitive to the global shape's orientation and investigate whether attentional capture driven by top-down attentional control settings for shapes are differently affected dependent on whether shapes have an ambiguous or unambiguous axis of symmetry.

Top-down Guidance of Visual Attention by More Complex Shapes

As we have only used simple shapes in our study, our results cannot be generalized for more complex shapes. However, certain aspects must be considered when the complexity of shapes is increased. For example, increasing complexity may hinder finding out which aspects of shapes guide visual attention, as this issue is still not solved even for simple shapes (Wolfe & Utochkin, 2019). Additionally, using more complex shapes may lead to a greater task difficulty due to a more complicated differentiation between target and distractors because the perceived similarity between shapes may increase the more local features (such as contour lines) they share. As the critical attributes accounting for shape similarity are still elusive, similarity effects occurring with different target-distractor combinations in visual search tasks are difficult to predict (Donnelly et al., 2000). In our study, the surprising effect in the diamond target condition may be driven by similarity based on shared contour lines, which arises the question of how top-down attentional control settings are affected by shared local features. However, research with shapes that share local features may provide further insight into the computation of similarity, and which aspects of shapes determine the guiding process of visual attention. Furthermore, research with complex shapes is of practical relevance as most of the natural shapes in our environment consist of multiple local features. Therefore, it is interesting how more complex shapes guide visual attention.

Conclusion

The present results shed further light on the role of shapes in the guiding process of visual attention and extend the current state of knowledge on whether shapes can be preattentively processed. We showed that specific shapes capture attention top-down based on search goals. Only matching cues produced a significant cueing effect, whereas non-matching cues did not control attentional allocation. Analyzing each target shape separately revealed that different shapes guide visual attention similarly efficient. However, the surprising effect in the diamond target condition indicates that top-down attentional control settings for shapes are not only activated by a specific shape but also by similar shapes that share critical attributes with the target. However, due to the small sample size and hence the low statistical power, the reliability of this effect must be tested in future studies, which should furthermore consider factors contributing to perceived shape similarity as they may moderate the effects under investigation.

Regarding the contingency of cueing effects on search goals for specific shapes, our results indicate that shapes can be at least to some extent preattentively processed. As the non-matching cue did not produce a significant cueing effect, it must have been processed and identified as not sharing a critical feature with the target before attention was allocated to its position. Our results provide several interesting avenues for future research that may investigate, for example, top-down guidance of attention by more complex shapes and whether top-down attentional control settings are sensitive to various external transformations such as orientational changes.

References

Adamo, M., Wozny, S., Pratt, J., & Ferber, S. (2010). Parallel, independent attentional control settings for colors and shapes. *Attention, Perception & Psychophysics*, *72*(7), 1730–1735. https://doi.org/10.3758/APP.72.7.1730

Ansorge, U., & Horstmann, G. (2007). Preemptive control of attentional capture by colour: Evidence from trial-by-trial analyses and orderings of onsets of capture effects in reaction time distributions. *Quarterly Journal of Experimental Psychology*, *60*(7), 952–975. https://doi.org/10.1080/17470210600822795

Ansorge, U., Kiss, M., Worschech, F., & Eimer, M. (2011). The initial stage of visual selection is controlled by top-down task set: New ERP evidence. *Attention, Perception, & Psychophysics*, *73*(1), 113–122. https://doi.org/10.3758/s13414-010-0008-3

Attneave, F. (1968). Triangles as ambiguous figures. *The American Journal of Psychology*, *81*(3), 447. https://doi.org/10.2307/1420645

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. Retrieved from https://github.com/crsh/papaja

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Belopolsky, A. V., Schreij, D., & Theeuwes, J. (2010). What is top-down about contingent capture? *Attention, Perception, & Psychophysics*, *72*(2), 326–341. https://doi.org/10.3758/APP.72.2.326

Berggren, N., & Eimer, M. (2018). Object-based target templates guide attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1368–1382. https://doi.org/10.1037/xhp0000541

Biderman, D., Biderman, N., Zivony, A., & Lamy, D. (2017). Contingent capture is weakened in search for multiple features from different dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(12), 1974–1992. https://doi.org/10.1037/xhp0000422

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Büsel, C., Voracek, M., & Ansorge, U. (2018). A meta-analysis of contingent-capture effects. *Psychological Research*. https://doi.org/10.1007/s00426-018-1087-3

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. https://doi.org/10.1016/j.visres.2011.04.012

Chen, P., & Mordkoff, J. T. (2007). Contingent capture at a very short SOA: Evidence against rapid disengagement. *Visual Cognition*, *15*(6), 637–646. https://doi.org/10.1080/13506280701317968

Donnelly, N., Found, A., & Müller, H. J. (2000). Are shape differences detected in early vision? *Visual Cognition*, *7*(6), 719–741. https://doi.org/10.1080/135062800501444407

Dowle, M., & Srinivasan, A. (2019). *Data.table: Extension of 'data.frame'*. Retrieved from https://CRAN.R-project.org/package=data.table

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *20*(8), 1423–1433. https://doi.org/10.1162/jocn.2008.20099

Eimer, M., & Kiss, M. (2010). Top-down search strategies determine attentional capture in visual search: Behavioral and electrophysiological evidence. *Attention, Perception, & Psychophysics*, *72*(4), 951–962. https://doi.org/10.3758/APP.72.4.951

Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 847–858. https://doi.org/10.1037/0096-1523.24.3.847

Folk, C. L., & Remington, R. W. (2008). Bottom-up priming of top-down attentional control settings. *Visual Cognition*, *16*(2), 215–231. https://doi.org/10.1080/13506280701458804

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 and Performance*, *18*(4), 1030–1044.

Found, A., & Müller, H. (1997). Local and global orientation in visual search. *Perception & Psychophysics*, *59*(6), 941–963.

Goller, F., & Ansorge, U. (2015). There is more to trial history than priming in attentional capture experiments. *Attention, Perception, & Psychophysics*, *77*(5), 1574–1584. https://doi.org/10.3758/s13414-015-0896-3

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480–506. https://doi.org/10.1037/0096-3445.121.4.480

Helmholtz, H. von. (1894). Über den Ursprung der richtigen Deutung unserer Sinneseindrücke [the origin and correct interpretation of our sense impressions]. *Zeitschrift Für Psychologie Und Physiologie Der Sinnesorgane*, *7*, 81–96.

Humphreys, G. W. (1983). Reference frames and shape perception. *Cognitive Psychology*, *15*(2), 151–196. https://doi.org/10.1016/0010-0285(83)90008-7

Humphreys, G. W., & Quinlan, P. T. (1988). Priming effects between two-dimensional shapes. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 203–220. https://doi.org/10.1037/0096-1523.14.2.203

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259. https://doi.org/10.1109/34.730558

Jongen, E. M. M., & Smulders, F. T. Y. (2007). Sequence effects in a spatial cueing task: Endogenous orienting is sensitive to orienting in the preceding trial. *Psychological Research*, *71*(5), 516–523. https://doi.org/10.1007/s00426-006-0065-3

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2010). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2010.00238

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lamy, D. F., & Kristjansson, A. (2013). Is goal-directed attentional guidance just intertrial priming? A review. *Journal of Vision*, *13*(3), 14–14. https://doi.org/10.1167/13.3.14

Lenth, R. (2019). *Emmeans: Estimated marginal means, aka least-squares means*. Retrieved from https://CRAN.R-project.org/package=emmeans

McCants, C. W., Berggren, N., & Eimer, M. (2018). The guidance of visual search by shape features and shape configurations. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(7), 1072–1085. https://doi.org/10.1037/xhp0000514

Palmer, S. E. (1985). The role of symmetry in shape perception. *Acta Psychologica*, *59*(1), 67–90. https://doi.org/10.1016/0001-6918(85)90042-3

Pinsk, M. A., Doniger, G. M., & Kastner, S. (2004). Push-pull mechanism of selective attention in human extrastriate cortex. *Journal of Neurophysiology*, *92*(1), 622–629. https://doi.org/10.1152/jn.00974.2003

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., & Snyder, C., R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55–85). Hillsdale, NJ: Erlbaum.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Theeuwes, J. (2013). Feature-based attention: It is all bottom-up priming. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1628), 20130055. https://doi.org/10.1098/rstb.2013.0055

Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). On the time course of top-down and bottom-up control of visual attention. In *Control of cognitive processes: Attention and performance XVIII* (pp. 105–125). Cambridge: MIT Press.

Torchiano, M. (2019). *Effsize: Efficient effect size computation*. https://doi.org/10.5281/zenodo.1480624

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research*, *37*(1), 25–43. https://doi.org/10.1016/S0042-6989(96)00111-3

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 0058. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., & Utochkin, I. S. (2019). What is a preattentive feature? *Current Opinion in Psychology*, *29*, 19–26. https://doi.org/10.1016/j.copsyc.2018.11.005

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science*, *2*(5), 156–161. https://doi.org/10.1111/1467-8721.ep10768973

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 121–134. https://doi.org/10.1037/0096-1523.16.1.121

List of Figures

List of Tables

Appendix

Abstract

We investigated whether specific shapes (triangle pointing downwards, triangle pointing upwards, square, and diamond) can guide visual attention in a top-down manner using a contingent capture paradigm. Our results indicate that top-down attentional control settings for specific shapes can be established. Reaction times were significantly faster in trials when the target was presented at the same position as the preceding matching cue (valid trials) compared to when the matching cue preceded the target at a different position (invalid trials). In contrast, reaction times were not significantly different in valid and invalid trials with non-matching cues. This finding indicates that attention was only captured by cues that shared the target feature. The results of analyzing each target shape condition separately corresponded to our main analysis (which included the data of all four target shapes) except for the diamond target condition. While we found pronounced contingent capture effects in all other target shape conditions, the cueing effect difference between matching and non-matching cues was scarcely evident in the diamond target condition. As it was already shown that similar stimuli capture attention as well, this surprising result indicates that the diamond is similar to the hexagon (both shapes share four similar oblique lines). There was no speed-accuracy trade-off found in the error rates.

Zusammenfassung

Wir untersuchten, ob einfache geometrische Formen (Dreieck, mit der Spitze nach unten gerichtetes Dreieck, Quadrat und Raute) die visuelle Aufmerksamkeit in Abhängigkeit eines Zielreizes steuern können. Wir fanden einen signifikanten Hinweisreiz-Effekt für passende Hinweisreize (die Hinweisreiz- und die Zielreizform waren identisch) und keinen signifikanten Effekt für nicht passende Hinweisreize (in jeder Zielreizform-Bedingung ausschließlich das Hexagon). In Durchgängen mit passenden Hinweisreizen waren die Reaktionszeiten in validen Durchgängen, wenn der Hinweisreiz an derselben Position wie der nachfolgende Zielreiz präsentiert wurde, schneller als in invaliden Durchgängen, wenn der Zielreiz an einer anderen Position auftrat als der zuvor präsentierte Hinweisreiz. In Durchgängen mit nicht passenden Hinweisreizen unterschieden sich die Reaktionszeiten nicht in Abhängigkeit von der Validität des Zielreizes. Die Ergebnisse der vier einzelnen Zielreizform-Analysen entsprachen weitgehend den Ergebnissen der allgemeinen Datenanalyse mit Ausnahme der Rauten-Bedingung, in der wir einen abweichenden Hinweisreiz-Effekt fanden. In dieser Zielreizform-Bedingung gab es keinen eindeutigen Unterschied im Hinweisreiz-Effekt zwischen passenden und nicht-passenden Hinweisreizen, der im Gegensatz dazu in den drei anderen Zielreizform-Bedingungen stark ausgeprägt war. Eine mögliche Erklärung für diesen Effekt ist, dass eine Raute vier diagonale Linien mit einem Hexagon teilt, was darauf hindeutet, dass auch ähnliche Hinweisreize Aufmerksamkeit anziehen. In den Fehlerraten zeigten sich keine signifikanten Unterschiede in den verschiedenen Zielreizform-Bedingungen.