Cognition 150 (2016) 150–162

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

# Inhibition drives configural superiority of illusory Gestalt: Combined behavioral and drift–diffusion model evidence

Qi-Yang Nie [a,*], Mara Maurer [a], Hermann J. Müller [a,b], Markus Conci [a]

[a] Department Psychologie, Ludwig-Maximilians-Universität München, Munich, Germany
[b] Department of Psychological Sciences, Birkbeck College, University of London, London, UK

ARTICLE INFO

*Article history:*
Received 6 July 2015
Revised 8 February 2016
Accepted 11 February 2016

*Keywords:*
Visual search
Kanizsa figure
Configural superiority
Inhibition
Drift–diffusion model

ABSTRACT

Illusory Kanizsa figures demonstrate that a perceptually completed whole is *more* than the sum of its composite parts. In the current study, we explored part/whole relationships in object completion using the configural superiority effect (CSE) with illusory figures (Pomerantz & Portillo, 2011). In particular, we investigated to which extent the CSE is modulated by closure in target and distractor configurations. Our results demonstrated a typical CSE, with detection of a configural whole being more efficient than the detection of a corresponding part-level target. Moreover, the CSE was more pronounced when grouped objects were presented in distractors rather than in the target. A follow-up experiment systematically manipulated closure in whole target or, respectively, distractor configurations. The results revealed the effect of closure to be again stronger in distractor, rather than in target configurations, suggesting that closure primarily affects the inhibition of distractors, and to a lesser extent the selection of the target. In addition, a drift–diffusion model analysis of our data revealed that efficient distractor inhibition expedites the rate of evidence accumulation, with closure in distractors particularly speeding the drift toward the decision boundary. In sum, our findings demonstrate that the CSE in Kanizsa figures derives primarily from the inhibition of closed distractor objects, rather than being driven by a conspicuous target configuration. Altogether, these results support a fundamental role of inhibition in driving configural superiority effects in visual search.



## 1. Introduction

> Whilst part of what we perceive comes through our sense from the object before us, another part (and it may be the larger part) always comes out of our own head.
>
> [James (1890, p. 103)]

Understanding how the retinal images of our complex visual world are translated into integrated and coherent object representations was recognized as a central challenge by Gestalt theory (Wertheimer, 1912). A major question in this context is how the visual system combines fragments into wholes despite adverse luminance gradients and partial occlusions of the underlying scene structure. Solving this problem, by means of perceptual organization, is a fundamental function of the visual system. A number of 'laws' have been proposed describing the organizational ('grouping') principles based on which the visual system structures our environment, including grouping based on proximity, closure, and symmetry (Wagemans, Elder, et al., 2012).

Empirical research has shown that the laws of grouping as described initially on the basis of phenomenological observations are essential for object recognition (Lowe, 1987). For example, parsing retinal images through mechanisms of perceptual organization may result in ordered scene representations where fragments are assigned unambiguously to a given object and each object can be segregated from elements that belong to other objects and the background. Such structured representations are achieved even when distinctive and continuous borders between objects are lacking. For instance, 'Kanizsa figures', such as the Kanizsa square depicted in Fig. 1C, demonstrate that mechanisms of visual completion can give rise to the impression of an illusory object – that is, in the example, a relatively bright central square with sharp boundaries emerges that appears to occlude the (adjacent) circular inducer elements (Kanizsa, 1955) – even though this percept has no direct physical correspondence in the retinal image (Murray & Herrmann, 2013, for a review). Original Gestalt theory claimed that closure, rather than just being a cue for grouping, is

* Corresponding author at: Allgemeine und Experimentelle Psychologie, Department Psychologie, Ludwig-Maximilians Universität, Leopoldstr. 13, D-80802 München, Germany.

*E-mail address:* qiyang.nie@psy.lmu.de (Q.-Y. Nie).

http://dx.doi.org/10.1016/j.cognition.2016.02.007
0010-0277/

a major determinant of what constitutes a complete form (Koffka, 1935). More recently, Elder and Zucker (1993, 1994) proposed that the most important role of closure may be to relate a 1-D contour to a corresponding 2-D shape – which was supported by the finding that small changes in closure can yield large changes in shape discriminability. In this view, emergent properties of illusory figures may reflect the degree to which grouping by closure yields a global form (Kogo, Strecha, Van Gool, & Wagemans, 2010; Kogo & Wagemans, 2013; Wagemans, Elder, et al., 2012). It should be noted that in the example of the Kanizsa square, the closed shape is not part of the actual (physical) stimulus arrangement, but is rather attributed to the emergent, illusory square – that is, it actually constitutes some form of "implied closure". Fig. 1 illustrates that implied closure of the emergent figure can be varied systematically by changing the configuration of the pacman inducers. Moreover, along with an increase in closure (from Fig. 1A to C), the emergent shape exhibits a concurrent increase in the extent to which precise bounding contours are perceived based on grouping by collinearity.

Despite the complex percepts that arise from illusory figures, arguably, such integrated objects are nevertheless rendered by preattentive coding mechanisms (Davis & Driver, 1994; see also Gurnsey, Humphrey, & Kapitan, 1992). For instance, Davis and Driver (1994) used a visual search task with a Kanizsa square as the target and comparable configurations (that did not give rise to an illusory figure) as nontargets. Davis and Driver found that search for an illusory target figure could be performed 'efficiently', that is, the reaction times (RTs) taken to respond to the presence of the target were independent of the number of configurations presented in the search display (the 'display size'). Subsequent studies, by Conci, Gramann, Müller, and Elliott (2006), Conci, Müller, and Elliott (2007a, 2007b) and Conci, Töllner, Leszczynski, and Müller (2011), showed efficient search for illusory figures to primarily rely on grouping by closure; that is, search efficiency, reflecting how readily focal attention is allocated to the target, was primarily determined by the degree of closure provided in the target and distractor configurations. By contrast, search efficiency was found to be unrelated to the contour information, that is, the degree to which emergent shapes are constructed on the basis of grouping by collinearity (Conci et al., 2006, 2007a, 2009; Donnelly, Humphreys, & Riddoch, 1991). Thus, converging evidence from studies that employed Kanizsa-type stimuli suggests that closed object configurations are particularly effective in guiding search at preattentive stages of processing (Conci et al., 2011, 2009; Stanley & Rubin, 2005).

A related paradigm designed to examine the effectiveness of the emergent properties of grouping was introduced by Eidels, Townsend, and Pomerantz (2008), Pomerantz (2003), Pomerantz and Portillo (2011), Pomerantz and Pristach (1989), Pomerantz, Sager, and Stoever (1977) and Wagemans, Feldman, et al. (2012). Their 'Configural Superiority Effect' (CSE) typically shows that RTs to localize a target among distractors can be significantly faster when 'irrelevant' context parts are added to an item so as to elicit the percept of a complete figure. Fig. 2A illustrates a schematic example of the odd-quadrant task typically employed to investigate the CSE. Participants are asked to determine which one of four presented elements is different (e.g., element B) from the other, homogenous distractors (e.g., element A). Then, an additional, 'task-irrelevant' context item (e.g., element C) is added to all objects, now producing novel stimulus pairs (e.g., BC and AC). While this irrelevant context C does not convey any task-relevant information per se, in certain cases, stimuli will group together to form a perceptual 'Gestalt' – with one such configuration providing salient information as to what constitutes the target, thus producing a CSE (see Fig. 2B for a typical example). For such configurations, detection (and localization) of the novel, composite

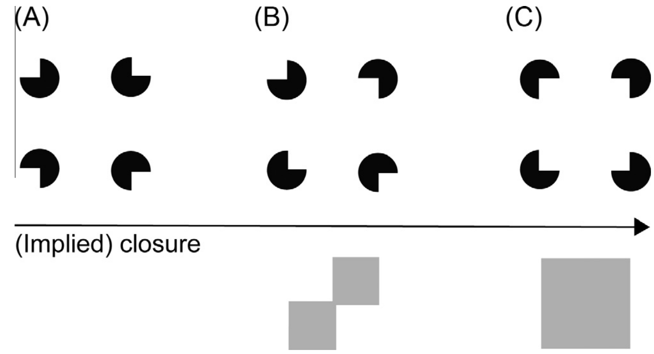


**Fig. 1.** Implied closure in emergent shape configurations. Panels A to C illustrate that a systematic (i.e., inward facing) arrangement of pacman inducers can modify the amount of closure in the emergent (illusory) figure until a 'complete' Kanizsa square (C) is rendered. Each stimulus configuration shows an arrangement of inducers (top) together with a schematic illustration of the corresponding emergent shape representation (bottom).

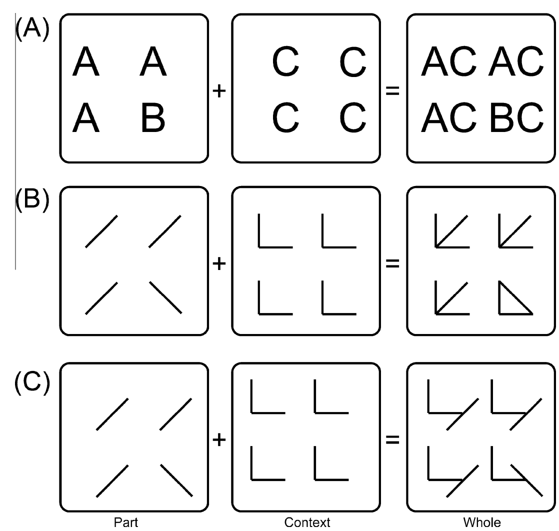


**Fig. 2.** Panel A shows a schematic of the odd-quadrant discrimination task. Participants see either a part or a whole display (the context display is only presented for illustrative purpose). In the example, A, B, and C are placeholders for several possible stimuli. Panel B shows a corresponding example of a stimulus set that leads to a reliable configural superiority effect, whereas panel C illustrates an example that yields a configural inferiority effect (adapted from Pomerantz & Portillo, 2011).

target becomes significantly easier (relative to the non-composite target), as evidenced by faster RTs and increased accuracy.

The CSE has been used to illustrate the major role of perceptual grouping for the extraction of basic 'Gestalts', or emergent features (Pomerantz & Portillo, 2011). CSEs have been reported for a variety of stimulus configurations. In one prototypical case, additional pacman inducers were presented that, in this variant, combined to form a non-square whole (target) among Kanizsa square (distractor) configurations, relative to a part condition that presented incomplete objects consisting of only two pacman inducers (see Fig. 3B). In general agreement with the findings from visual search paradigms (Davis & Driver, 1994), presentation of whole Kanizsa figures led to a reliable CSE. In many other cases, though, adding contextual information dilutes the differences between the two elements A and B, making it harder to discern the presence of the composite stimulus BC among stimuli AC, as compared to discerning stimulus B among stimuli A alone. Moreover, adding a context may also increase the total processing load, as well as

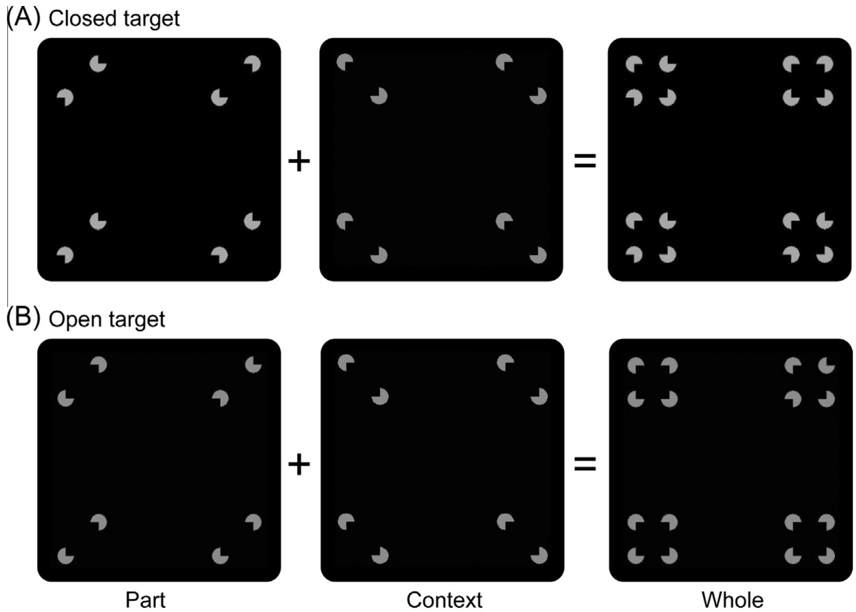


**Fig. 3.** Example search displays in Experiment 1. (A) In the closed target condition, a closed target was presented among open distractors. (B) In the open target condition, the assignment of targets and distractors was reversed. Both open and closed targets conditions were presented either as part or as whole displays. Whole displays combined the part display with a non-informative context display, to reveal complete configurations that typically yield a configural superiority effect.

increasing the chances of 'crowding', or observers may tend to attend to the wrong element (Pomerantz et al., 1977). This is referred to as 'Configural Inferiority Effect' (CIE; see Fig. 2C for an example), because the composite (whole) is significantly harder to discriminate than the corresponding part elements.

Consistent with the behavioral evidence on the CSE, a recent fMRI study suggests that the ventral visual pathway, in particular the Lateral Occipital Complex (LOC), is involved in the configural processing of emergent features (Kubilius, Wagemans, & Op de Beeck, 2011). Using a localization task (see Fig. 2B), this study showed that decoding of neuronal responses in LOC, but not in the primary visual cortex (V1), was better able to predict the location of the odd item when processing wholes, whereas area V1 (but not LOC) was a better predictor of the position of the odd element when processing parts. This pattern supports the idea that Gestalt configurations may emerge at a relatively higher level of visual processing, with the processing of parts and wholes being related to distinct areas, or stages, in the visual processing hierarchy.

The aim of the present study was to further explore the crucial processes that determine the CSE. For instance, reliable CSEs have been reported for a variety of stimulus configurations, thus providing evidence for the idea that perceptual grouping generates emergent features that allow for an efficient extraction of a given target configuration. However, these studies have – to our knowledge – not investigated in detail whether the detection of a configural target is enhanced because of emergent properties of the target (thus facilitating target detection), or due to emergent features in distractors (i.e., permitting more efficient distractor suppression). On the basis of these considerations, we set out to specifically test and compare how grouping in targets and distractors modulates the CSE.

To this end, Experiment 1 employed a variant of a CSE paradigm presenting circular pacman inducer elements that potentially combine to form an illusory Kanizsa figure (i.e., Pomerantz & Portillo, 2011). The experiment consisted of two task sessions: observers were required to detect either a closed target among open nontargets (Fig. 3A) or an open target among closed nontargets (Fig. 3B). Importantly, the target could be presented within either a 'Part' or a 'Whole' display (Fig. 3, left and right panels, respectively). Comparisons of the two possible target configurations permit us to examine whether the CSE with illusory figures can be related to grouping by closure in targets and/or distractors (Fig. 3A and B, respectively). Next, to further disentangle grouping by closure in either targets or distractors, Experiment 2 introduced separate experimental parts that independently manipulated the degree of closure in distractors or, respectively, in the target (while keeping the target or, respectively, the distractors constant, see Fig. 7). This approach allowed examination for the separate, independent contributions of closed configurations in targets and distractors.

Moreover, while previous behavioral studies reported reliable RT effects, it is not clear at which functional level of processing the CSE emerges – that is, whether the CSE can be related to basic levels of information processing or to higher-level, decisional stages. For instance, CSE differences across conditions may reflect differences in the rate at which stimulus information is accumulated (the so-called 'drift rate'), the amount of decisional information required to provide a response (i.e., 'boundary separation'), or other nondecisional factors that influence the response, in particular initial sensory processing ('non-decision time'; see Ratcliff & McKoon, 2008). To our knowledge, there have been no attempts to model perceptual grouping by means of such a diffusion-type modeling approach. Thus, to examine such latent processing stages, we applied a model fitting procedure to the behavioral data using the Hierarchical Drift–Diffusion Model (HDDM; Wiecki, Sofer, & Frank, 2013), which incorporates an estimation of these parameters, in addition to the conventional response latency and accuracy measures.

To preview our main findings, both experiments consistently revealed that the CSE or search for a configural (Kanizsa-type) target are primarily determined by grouping by closure in distractors, but not in the target configuration. This suggests that closed shapes can be more readily rejected (as a result, the target is detected more efficiently). Our modeling results further reveal that this effect of closure in distractors is reflected in the drift rates, that is, faster rates of evidence accumulation to reach a given decision when distractor shapes are bound to a coherent (closed) object. In this view, the CSE is determined by the inhibition of to-be-rejected distractor configurations.

## 2. Experiment 1

Experiment 1 investigated object grouping, that is, grouping by closure in target and distractor configurations, using a visual search task with Kanizsa-type configurations (see Fig. 1 for examples, and Pomerantz & Portillo, 2011). The target configuration could be presented either as a whole or as a part configuration (see Fig. 3, left and right panels, respectively). Two conditions presented either a closed target among open distractors, or, conversely, an open target among closed distractors (Fig. 3, panels A and B, respectively). Differences between targets and distractors were kept constant across wholes and parts such that a given target would always yield an identical feature contrast value relative to the distractors.[1] On the basis of previous findings, we expected faster RTs to whole as compared to part configurations, which would be indicative of a CSE (Pomerantz & Portillo, 2011; Pomerantz et al., 1977).

### 2.1. Methods

#### 2.1.1. Participants

Fourteen right-handed observers (10 female; age range: 21–28 years; mean age: 23.6 years) with normal or corrected-to-normal visual acuity participated in the experiment, receiving course credits or payment of 8 Euro per hour. Participants provided written consent to the procedure of the experiment, which was approved by the ethics committee of the Department of Psychology at LMU München, in accordance with the Declaration of Helsinki.

#### 2.1.2. Apparatus and stimuli

The experiment was conducted on a PC-compatible computer (Dell Inc., Texas, USA) using Matlab routines and Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli were presented on a 22″ LCD monitor screen placed at a viewing distance of approximately 57 cm. Stimuli were presented in light gray (8.5 cd/m$^2$) against a black (0.02 cd/m$^2$) background. On each trial, four configurations were placed within the quadrants of the screen, 6° from the screen center. Each configuration subtended 2.8° × 2.8° of visual angle and was composed of two or four pacman inducers, with a diameter of 0.93° each.

Configurations could be presented as 'parts' or 'wholes', presenting two or four pacman inducers, respectively. Part configurations presented two pacmen aligned along an imaginary diagonal line across the quadrant (see Fig. 3, left panels). Left- or rightward tilt of the diagonal was chosen randomly for a given trial, though with each display presenting all objects in uniform orientation (i.e., for a given display all distractors were identical). Whole configurations presented four pacmen arranged in square form (see Fig. 3, right panels). Configurations with pacman inducers rotated such that all indented segments faced the center of the configuration are referred to as 'closed'; by contrast, when two pacmen faced outwards, the resulting configurations are referred to as 'open'.

Within a given trial, only whole or only part configurations were displayed. Two types of trials were possible: On target-present trials, one target configuration was presented among three distractor configurations, with either a closed target among open distractors, or an open target among closed distractors. On target-absent trials, all four configurations on a given trial were identical. Fig. 3 presents examples of target-present displays. The figure illustrates how whole configurations were created by combining a given part display with an additional, 'uninformative' context display.

#### 2.1.3. Design

A three-factors within-subjects design was used. The independent variables were target presence, configuration, and target closure. Target presence had two levels: target present and target absent. For target-present trials, there was always one configuration that differed from the other three, whereas for target-absent trials, all four configurations were the same. Targets appeared with equal probability at the four possible display locations, with target location varying randomly across trials. The second variable, configuration, also had two levels: whole and part, denoting whether a given display consisted of stimulus arrangements made up of four or two pacman inducers, respectively (see Fig. 2 and the descriptions above for further details). The third variable, target closure, again had two levels: closed and open (see Fig. 3A and B, respectively), denoting whether a given target could be grouped to form a closed shape or not. Closed targets were presented with open distractors, and open targets with closed distractors.

#### 2.1.4. Procedure

Participants were comfortably seated in a dimly lit, sound-attenuated room. The experiment was divided into two consecutive sessions that either presented closed or open targets (with order of presentation counterbalanced across observers). Each session started with 48 practice trials for participants to become familiar with the task. Then, in each session, 256 experimental trials were presented in four blocks of 64 trials each, with randomized order of the factors target presence and configuration. There were 64 trials for each factorial combination.

Each trial started with the presentation of a central fixation cross for 500 ms. Subsequently, a search display was presented until the observer's response. Participants responded with a speeded target-present versus target-absent response via mouse keys.[2] The response mapping (i.e., left/right-hand responses to target presence/absence) was counterbalanced across participants. In case of an erroneous response, feedback was provided by an alerting message (a red minus sign) that was presented for 1000 ms in the center of the screen. Each trial was separated from the next by an interval of 500 ms, presenting a blank screen.

### 2.2. Results

#### 2.2.1. Response accuracy

Overall, performance was very accurate, with an average of 94% correct responses. Fig. 4A depicts the accuracy data (percentage of correct responses), which were examined by a 2 × 2 × 2 repeated-measures analysis of variance (ANOVA) with the factors target presence (present, absent), configuration (whole, part), and target closure (closed, open). We additionally report the estimated Bayes factor (BF) for all significant results, as revealed by a comparable Bayesian ANOVA using JASP (Love et al., 2015). The Bayes factor gives the ratio with which the alternative hypothesis is favored over the null hypothesis (i.e., larger BFs argue in favor of the alternative hypothesis; see Dienes, 2011, for an overview). The accuracy ANOVA revealed the main effects of both configuration (wholes vs. parts: 95% vs. 92%, $F(1,13) = 5.38$, $p = .037$, $\eta^2 = .29$, BF = 3.09) and target closure (closed vs. open: 92% vs. 95%, $F(1,13) = 18.64$,

[1] The stimulus set used in both experiments was carefully controlled in terms of the similarity relations between target and distractors. Nevertheless, it remains possible that some subjective components of similarity (cf., Hout et al., 2016) were not captured by our control of the stimulus parameters.

[2] It should be noted that CSE tasks usually employ a quadrant localization task (Pomerantz et al., 1977) whereas here we used a detection task. This slight change of the paradigm was implemented in order to apply diffusion modeling to the data (which requires two response alternatives). However, both types of task are usually highly comparable (e.g., Green, 1992).

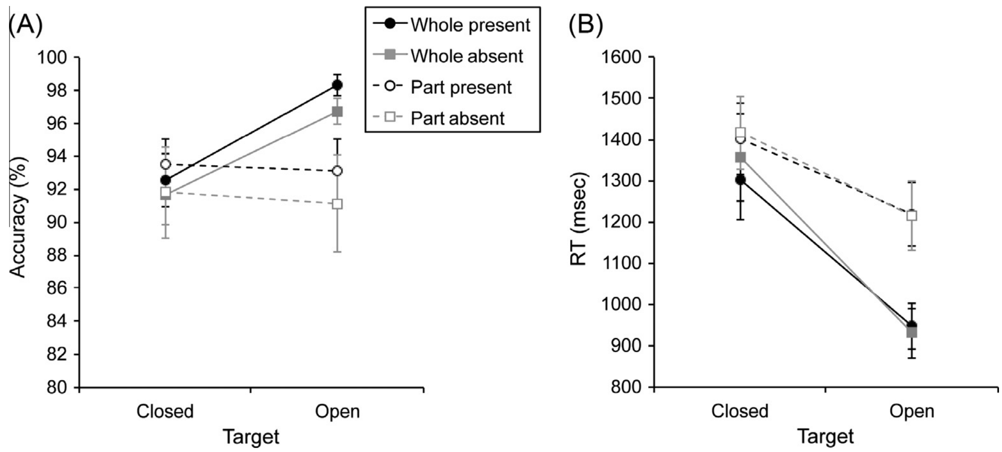


**Fig. 4.** Mean accuracy (A) and mean RTs (B) in Experiment 1 presented as a function of target closure (closed vs. open) for the factorial combinations of configuration (whole, part) and target presence (present, absent). The error bars represent ±1 standard error of the mean.

$p < .001$, $\eta^2 = .59$, BF = 3.22) to be significant. Importantly, the two-way interaction between configuration and target closure was also significant, $F(1,13) = 11.26$, $p = .005$, $\eta^2 = .46$, BF = 5.51. Post-hoc comparisons revealed a CSE in accuracy: there was a reliable difference in response accuracy only for open targets (5.5%; $t(13) = 3.8$, $p = .002$, $d = 1.01$, BF = 19.7), but not for the closed targets (−0.3%, $t(13) = -.47$, $p = .65$, $d = -0.13$, BF = 0.3). Neither the main effect nor any interactions involving the factor target presence were significant (all *ps* > .4, $\eta^2$s < .05, BFs < 0.3). This pattern of results suggests that a CSE in accuracy was evident only for open targets (among closed distractors), without a comparable facilitatory effect for closed targets (among open distractors). Moreover, the CSE in accuracy was found to be independent of target presence.

### 2.2.2. Reaction times

Mean RTs for each observer were calculated excluding error responses and RTs deviating by more than three standard deviations from the mean of each condition. 7.7% of all trials, on average, were excluded by this outlier criterion (Experiment 2 yielded comparable exclusion rates). Mean RTs were again entered in a $2 \times 2 \times 2$ repeated-measures ANOVA with the factors target presence (present, absent), configuration (whole, part), and target closure (closed, open). Fig. 4B depicts the RT results. The analysis revealed both the main effect of configuration (wholes vs. parts: 1119 vs. 1325 ms, $F(1,13) = 21.82$, $p < .001$, $\eta^2 = .62$, BF = $3.84 \times 10^{10}$) and that of target closure (closed vs. open: 1371 vs. 1078 ms, $F(1,13) = 48.4$, $p < .001$, $\eta^2 = .79$, BF = $3.15 \times 10^4$) to be significant. Moreover, a significant interaction between target closure and configuration was again found ($F(1,13) = 13.05$, $p = .003$, $\eta^2 = .5$, BF = 20.28). This interaction was owing to a reliable CSE, of 285 ms, for open targets (presented among closed distractors), $t(13) = -7.65$, $p < .001$, $d = -2.05$, BF = 5431.4. By contrast, for closed targets (presented among open distractors), the CSE (of 108 ms) was not significant, $t(13) = -1.48$, $p = .16$, $d = -0.4$, BF = 0.66. Again, there was no main or interaction effect that involved target presence (all *ps* > .48, $\eta^2$s < .04, BFs < 0.16), mirroring the pattern in the accuracy data. This pattern of results shows, as above, that the CSE was particularly pronounced for closed distractors, without any substantial contribution arising from target presence and/or target closure.

## 2.3. Hierarchical drift–diffusion modeling

In a subsequent step, a drift–diffusion modeling approach was applied to further demarcate task-critical stages determining the CSE. We used the Hierarchical Drift–Diffusion Modeling approach (HDDM; Wiecki et al., 2013) to (i) apply a model fitting procedure and (ii) extract model parameters of the best-fitting model for further analysis. Theoretically, the diffusion model specifies decision processes with two possible outcomes (e.g., deciding between target presence and absence) as being inherently noisy, with information being accumulated over time. It permits the extraction of three parameters relating to, respectively: (1) information accumulation, which can be interpreted as a general measure of sensivity to the relevant configurations (the 'drift rate' parameter, $v$); (2) a decision threshold reflecting the amount of information required to trigger the corresponding response (the 'boundary separation' parameter, $a$); and (3) a mean 'non-decision' time parameter ($T_{er}$), which refers to the time taken by the sensory encoding of the information plus the time required for executing the motor response (Ratcliff & McKoon, 2008). It should be noted that motor responses can be assumed to reflect a constant process on all types of trials (as they are issued on every single trial); accordingly, potential differences in non-decision times could be taken to reflect exclusively the stage(s) of initial sensory processing.

HDDM constitutes a recently developed hierarchical Bayesian estimation of drift–diffusion parameters based on the RT distributions of both correct and incorrect responses, allowing for a simultaneous extraction of individual and group parameters. Fits to individual participants are constrained by the group distribution but can deviate from this distribution to a certain extent reflecting individual variability. To compare choice RTs in the CSE, eight different models were investigated, where the three parameters of interest ($v$, $a$, $T_{er}$) were either fixed or allowed to vary across the eight model variants (Table 1). For each model, there were 20,000 samples generated from the posterior probabilities, where the first 2000 samples were discarded. Of the remaining 18,000 samples, every fifth sample was saved, resulting in a trace of 3600 samples. The best model to describe the data across the eight conditions was selected on the basis of the deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002), reflecting the best trade-off between the quality of fit and model complexity. To evaluate model performance, posterior predictives generated by the winning model were plotted on top of the observed correct and incorrect RT distributions for each participant. Fig. 5 represents an example of one representative participant.

As depicted in Table 1, this model selection procedure showed the best fit when all three parameters (drift rate $v$, boundary separation $a$, nondecision time $T_{er}$) were allowed to vary (model 1,

**Table 1**
Model selection with HDDM in Experiment 1. A lower value of the deviance information criterion (DIC) indicates a better balance between model fit and complexity. $v$ = drift rate; $a$ = boundary; $T_{er}$ = nondecision time.

| Model | Free to vary | DIC |
|---|---|---|
| **1** | $\boldsymbol{v, a, T_{er}}$ | **7074.9** |
| 2 | $v, T_{er}$ | 7308.8 |
| 3 | $v, a$ | 7422.5 |
| 4 | $a, T_{er}$ | 7555.0 |
| 5 | $T_{er}$ | 8009.6 |
| 6 | $a$ | 8104.4 |
| 7 | $v$ | 8281.5 |
| 8 | Fix all | 10766.3 |

printed in bold), corresponding to a full drift–diffusion model. Next, each parameter of this best fitting model was then entered into a 2 × 2 × 2 repeated-measures ANOVA with the factors target presence, configuration, and target closure, as for the above analyses.

First, analysis of the *drift rates* revealed significant main effects for configuration (wholes vs. parts: 2.15 vs. 1.71, $F(1,13) = 17.09$, $p = .001$, $\eta^2 = .57$, BF = $1.61 \times 10^{14}$) and target closure (closed vs. open: 1.61 vs. 2.27, $F(1,13) = 153.4$, $p < .001$, $\eta^2 = .92$, BF = $4.6 \times 10^9$). These main effects indicate that the rate of evidence accumulation was faster for wholes relative to parts, and for open relative to closed targets. There was also a configuration by target closure interaction ($F(1,13) = 82.21$, $p < .001$, $\eta^2 = .86$, BF = $2.78 \times 10^7$). Post-hoc paired $t$-tests showed the CSE to be significant only for open targets ($t(13) = 9.78$, $p < .001$, $d = 2.61$, BF = $6.42 \times 10^4$; whole vs. part: 2.77 vs. 1.77), but not for closed targets ($t(13) = -0.88$, $p = .4$, $d = -0.24$, BF = 0.38; whole vs. part: 1.53 vs. 1.65). As can be seen from Fig. 6A, a CSE in drift rates was evident in the open, but not in the closed, target condition.

Next, a repeated-measures ANOVA of the *decision thresholds* revealed only the interaction between target presence and target closure to be significant ($F(1,13) = 5.06$, $p = .042$, $\eta^2 = .28$, BF = 0.093). Post-hoc paired $t$-tests showed the main effect of target closure to be marginally significant for target absent trials ($t(13) = 1.78$, $p = .099$, $d = 0.48$, BF = 0.94; open vs. closed: 2.76 vs. 2.54), but not for target present trials ($t(13) = -1.25$, $p = .23$, $d = -0.33$, BF = 0.52). Thus, open distractor configurations tended to require more evidence to be accumulated than closed configurations in order to reach the target-absent decision boundary (Fig. 6B).

Finally, a repeated-measures ANOVA of the *nondecision times* yielded significant main effects of both configuration ($F(1,13) = 22.23$, $p < .001$, $\eta^2 = .63$, BF = $2.28 \times 10^4$) and target closure ($F(1,13) = 34.03$, $p < .001$, $\eta^2 = .72$, BF = $1.67 \times 10^5$). As can be seen from Fig. 6C, wholes were encoded faster than the corresponding parts (494 vs. 633 ms) and the same was true for closed distractors (open targets) versus open distractors (closed targets) (closed vs. open distractors: 500 vs. 627 ms); that is, sensory encoding of stimulus information was actually more efficient with both larger amounts of physical stimulation and with closed configurations. There were no further significant effects ($p$s > .23, $\eta^2$s < .11, BFs < 0.6).

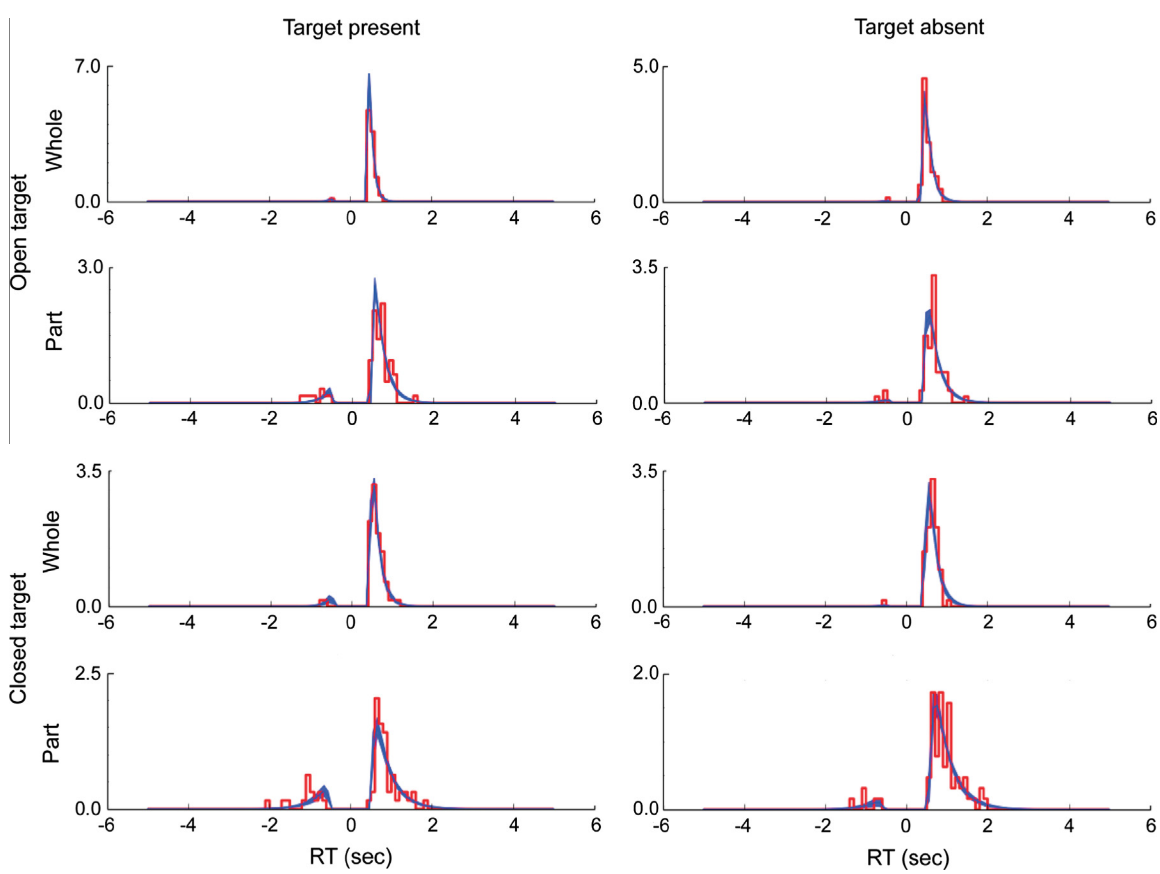


**Fig. 5.** Examples of the posterior predictive distribution as extracted from the optimal HDDM (blue lines), and the respective empirical normalized RT distributions from one representative participant in Experiment 1 (red lines). Each panel depicts the distributions for separate conditions in the experiment. Errors have been mirrored along the $x$-axis to display correct and incorrect RT distributions in one plot (positive and negative values, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

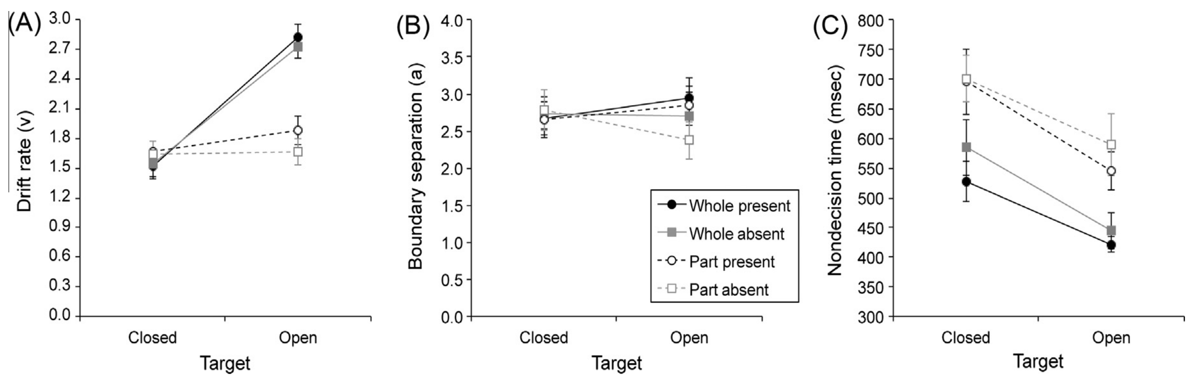


**Fig. 6.** Mean hierarchical drift–diffusion parameters (A: drift rate; B: boundary separation; C: nondecision time) in Experiment 1. All parameters are presented as a function of target closure (closed vs. open) for the factorial combinations of configuration (whole, part) and target presence (present, absent). The error bars represent ±1 standard error of the mean.

### 2.4. Discussion

The results of Experiment 1 replicated previous findings of a CSE with illusory figures (Pomerantz & Portillo, 2011; Pomerantz et al., 1977). Overall, wholes were detected 180 ms faster than the corresponding parts, demonstrating that a given configuration can be processed faster than its constituent elements. Importantly, however, this overall pattern was differentially influenced by target closure: a much larger CSE manifested with an open target (presented among closed distractors), as compared to a closed target (among open distractors; CSEs of 285 [108] ms for open [closed] targets, respectively), indicating that the magnitude of the CSE is modulated by the degree of closure in distractors. Notably, the fact that a robust CSE was obtained only in the condition in which a closed Kanizsa square served as the distractor (but not when the target was a closed Kanizsa square, in which case the CSE was not reliable) would suggest that the emergence of the CSE is primarily associated with the suppression of (closed) distractors, rather than selection of a (closed) target. Moreover, target-absent trials were overall comparable to target present trials, further suggesting that closure is primarily modulating the efficient rejection of a given distractor configuration.

In addition, the drift–diffusion model analysis further identified specific processing stages associated with this CSE-related influence of distractors. The parameter estimates obtained indicate that the initial visual encoding processes, reflected by the *nondecision times*, were affected by object closure, illustrating that closed configurations were encoded more efficiently; however, they were also influenced by the amount of visual stimulation provided – as evidenced by the faster processing of wholes as compared to the corresponding parts (Fig. 6C). A difference between closed and open configurations in the CSE was revealed only for subsequent processing stages reflected in the *drift rates*, with faster rates of evidence accumulation for wholes, relative to parts, with open targets [and closed distractors], as compared to closed targets [and open distractors] (Fig. 6A). This pattern mirrors that of the CSE in the RT data (Fig. 4B), suggesting that efficient rejection of closed distractors can expedite the accumulation of decision-critical evidence in favor of target presence. Next, the analysis of the *decision thresholds* (Fig. 6B) revealed somewhat larger thresholds to reach an open (vs. closed) decision boundary on target-absent trials, but no such difference on target-present trials. In sum, the hierarchical drift diffusion modeling disclosed distinctive dynamics at different processing stages: initial stimulus encoding was expedited with both larger amounts of visual stimulation and closed configurations presented, whereas a difference that reflected the (differential) CSE in open and closed targets was manifest at the subsequent stage of evidence accumulation only. Finally, the decision threshold tended to be (marginally) higher for closed distractors (but only when there was no target).

## 3. Experiment 2

Experiment 1 revealed a CSE that was primarily related to the processing (i.e., to the rejection) of closed distractors, manifesting in terms of both expedited RTs and the speed of evidence accumulation. Experiment 2 was designed to systematically examine the independent contribution of grouping by closure to the two (related) processes of target detection and distractor rejection. To this end, in Experiment 2, we only presented complete (whole) stimulus configurations that varied with regard to the amount of closure in either targets or distractors. There were two task sessions: First, in the '*distractor rejection task*', the distractors could be closed or open configurations and the target was held constant, presenting a 'mixed' configuration that was equally similar to both types of distractors (see Fig. 7A). Second, in the '*target detection task*', the target could be either a closed or an open configuration, whereas distractors were constant, always presenting a mixed configuration (see Fig. 7B). Therefore, these two tasks permit us to quantify closure (closed vs. open) separately in targets and distractors, and further to differentiate its relative contributions to target detection and distractor rejection. On the basis of Experiment 1, we expected that this manipulation would engender a more robust 'closure effect' in distractors than in targets.

### 3.1. Methods

#### 3.1.1. Participants

Fourteen right-handed observers (10 female; age range: 20–32 years; mean age = 25.6 years) with normal or corrected-to-normal visual acuity participated in Experiment 2, receiving course credits or payment of 8 Euro per hour.

#### 3.1.2. Apparatus and stimuli

All methodological details were essentially the same as in Experiment 1, except that, in Experiment 2, only whole configurations were presented – in three variants: they were (i) arranged to form a closed shape (i.e., a Kanizsa square), or (ii) depicted a mixed arrangement (with two diagonally opposing pacmen facing inwards and the other two pacmen facing outwards), or (iii) could be presented to form an open, symmetric shape (with all four pacmen oriented outwards). Fig. 1 presents examples of the closed, mixed, and open configurations (see also Fig. 7 for example dis-

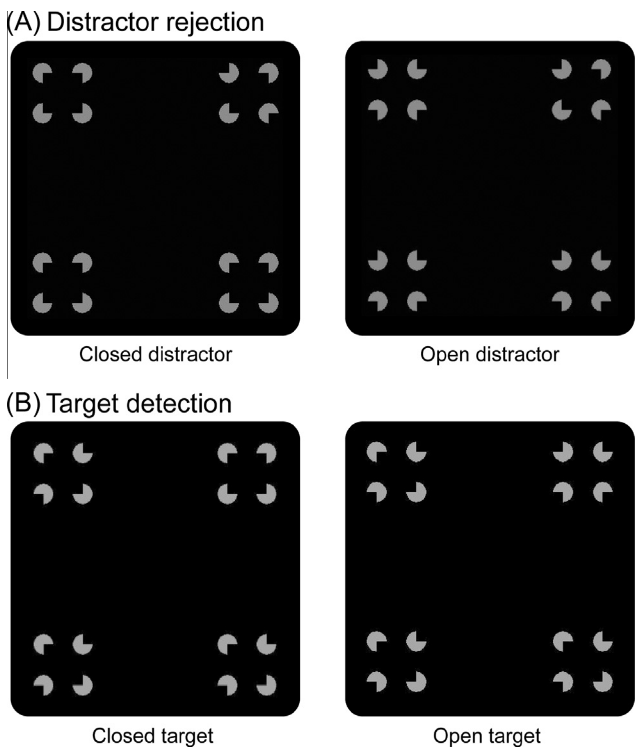


**Fig. 7.** Example search displays in Experiment 2. (A) In the distractor rejection task, closure in distractors was varied while keeping the target constant. (B) In the target detection task, distractors were constant but the target varied in terms of grouping by closure. Note that in all possible displays, the feature contrast between a given target and distractor configuration was the same, i.e., targets and distractors differed from each other to the same extent.

plays). As in Experiment 1, all distractors in a given search display were identical, homogeneous shapes. Note that the open configuration as previously used in Experiment 1 is now, in this variant of the task, referred to as 'mixed' configuration.

#### 3.1.3. Design and procedure

As in Experiment 1, the task in Experiment 2 was to detect a target that differed from the other configurations, and to respond with a speeded target-present or -absent response (with response mappings counterbalanced across observers).

The experiment consisted of two different halves, presented to observers in counterbalanced order: In one half of the experiment, the composition of the distractors was varied and the target remained constant throughout – so as to test the efficiency of rejecting closed or open distractors. Thus, in this part of the experiment, observers were required to detect a 'mixed' target among (variably across trials) either 'open' or 'closed' distractors. In the second half of the experiment, in turn, the target was varied and the distractors remained constant – to test the efficiency of detecting closed or open targets. This part of the experiment always presented 'mixed' configurations as distractors and observers were required to either detect an 'open' or a 'closed' target configuration. Fig. 7 presents examples of closed and open target-present trials for variations of both distractors (Fig. 7A) and targets (Fig. 7B). In each part of the experiment, target-present/-absent and closed/open configurations were presented in random order across trials. Targets were randomly assigned to one of the four display quadrants. There were 64 trials for each factorial combination. Each half of the experiment started with one practice block of 48 trials and was followed by 4 experimental blocks of 64 trials each.

### 3.2. Results

In order to directly compare the effects of grouping by closure in the processing of distractors and targets, the accuracies (and RTs) of target-present closed distractor conditions were subtracted from those in the corresponding open distractor conditions for each participant, thus providing a measure of the 'closure effect' in distractors (i.e., the benefit in accuracy and RTs for closed relative to open distractors). The same subtraction procedure was also applied to the closed and open target conditions. For statistical analysis, closure effects in targets and distractors were compared in a series of paired *t*-tests. Additional one-sample *t*-tests were employed to further investigate whether the obtained closure effects differed significantly from zero. Additional analyses of the mean target-present and -absent RTs and response accuracies (i.e., without applying this subtraction procedure) are presented in a Supplement.

#### 3.2.1. Response accuracy

A paired-sample *t*-test on the closure effect in the percentage of correct responses between the experimental halves related to distractor rejection and target detection, respectively, revealed no significant difference (6.7% vs. 5.8%, respectively; $t(13) = -0.21$, $p = .84$, $d = -0.06$, BF = 0.28; see Fig. 8A). Moreover, only the closure effect in distractors, but not that in targets, was significantly smaller than zero ($t(13) = 2.74$, $p = .009$, $d = 0.73$, BF = 7.28, and $t(13) = 1.53$, $p = .075$, $d = 0.41$, BF = 1.27, respectively), suggesting more accurate responses in rejecting closed than open distractor configurations, which is consistent with the pattern of the CSE in accuracy as obtained in Experiment 1.

#### 3.2.2. Reaction times

The same analysis procedure for the closure effect as above was applied. This analysis showed that the closure effect in distractors was significantly larger than that in targets (382 vs. 149 ms; $t(13) = 2.63$, $p = .02$, $d = 0.7$, BF = 3.13; see Fig. 8B), though the effects were significantly larger than zero in both cases ($ts(13) > 2.33$, $ps < .02$, $ds < 0.62$, BFs > 3.92). This indicates that closure facilitated both the detection of a (closed) target and the rejection of (closed) distractors, with closure in distractors yielding larger benefits for search performance – a finding again consistent with the results obtained in Experiment 1.

### 3.3. Hierarchical drift–diffusion modeling

As in Experiment 1, the HDDM modeling was applied to the data in order to identify the effect-critical stages of processing. The initial model-fitting procedure again supported a model variant where the three parameters (drift rate $v$, decision threshold $a$, nondecision time $T_{er}$) were all allowed to vary across conditions (see Table 2, model 1, printed in bold), thus, optimally predicting the observed RTs. Fig. 9 represents an example model fit for one representative participant. As for the RTs, to assess the magnitude of the closure effect, open minus closed distractor conditions (difference) scores for the various parameters as estimated by the best-fitting models were examined by statistical analyses.

First, the closure effect on the *drift rates* was computed. Note that for drift rates, more negative values correspond to a benefit for the closed configuration (whereas positive values would denote a cost), that is, the polarity of the effect is reversed for this diffusion parameter (relative to the pattern in RTs). A comparison of the drift rates between distractor and target processing revealed a significant difference ($-1.03$ vs. $-0.41$; $t(13) = -2.26$, $p = .04$, $d = -0.6$, BF = 1.81; see Fig. 10A), revealing a benefit of closure in the rate of evidence accumulation, which was particularly strong for distractor-related processing as compared to a weaker effect for

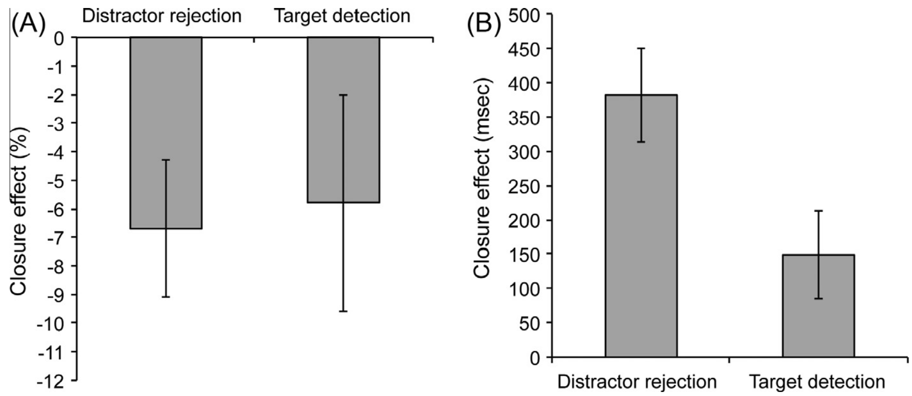


**Fig. 8.** Behavioral results from Experiment 2. (A) Mean closure effect in accuracy (mean accuracies for open minus closed configurations), and (B) mean RT closure effect (mean RTs for open minus closed configurations) for variations of the distractors and the target, respectively. The error bars represent ±1 standard error of the mean.

**Table 2**
Model selection with HDDM in Experiment 2.

| Model | Free to vary | DIC | |
|---|---|---|---|
| | | Distractor rejection | Target detection |
| **1** | **$v$, $a$, $T_{er}$** | **3581.7** | **6447.9** |
| 2 | $v$, $T_{er}$ | 3628.7 | 6628.2 |
| 3 | $v$, $a$ | 3730.9 | 6632.3 |
| 4 | $a$, $T_{er}$ | 3908.7 | 6673.6 |
| 5 | $T_{er}$ | 4083.0 | 6782.6 |
| 6 | $a$ | 4073.8 | 7124.1 |
| 7 | $v$ | 3895.2 | 7536.8 |
| 8 | Fix all | 4772.7 | 8177.6 |

target-related processing. In addition, both distractor- and target-related closure effects in drift rates showed a (marginally) significant difference from zero (distractor: $t(13) = -4.9$, $p < .001$, $d = -1.31$, BF = 224.5; target: $t(13) = -1.66$, $p = .06$, $d = -0.44$, BF = 1.52), indicating that the speed of evidence accumulation was overall faster for closed than for open configurations.

Next, the analysis of the closure effect on the *decision thresholds* revealed no significant results (all $ps > .2$, $ds < 0.22$, BFs < 0.36; Fig. 10B). This pattern indicates that the amount of decisional information required for rejecting closed distractors was comparable to that for rejecting open distractors.

Finally, the closure effect on *nondecision times* showed no significant difference between distractor- and target-related processing (132 vs. 149 ms, respectively; $t(13) = -0.32$, $p = .75$, $d = -0.09$, BF = 0.28; see Fig. 10C), suggesting that the duration of stimulus encoding was equivalent for comparisons of closure in distractor and target configurations. Both distractor- and target-related closure effects in non-decision times were significantly larger than zero ($ts(13) > 3.38$, $ps < .002$, $ds > 0.9$, BFs > 20.4), indicating that stimulus encoding of closed objects was more efficient than that of open configurations, irrespective of whether targets or distractors were varied.

### 3.4. Discussion

Experiment 2 revealed a more robust closure effect for distractors as compared to targets, replicating the general pattern of effects observed in Experiment 1. However, closure nevertheless also influenced the efficiency of target detection, albeit to a smaller extent. Overall, participants were more efficient both in rejecting closed distractors and in detecting the closed target, possibly because the Kanizsa square combines both closure and symmetry, while the open configuration is only symmetric but lacks closure and is, thus, less conspicuous.

The results of the hierarchical drift–diffusion modeling further demonstrated the underlying sources of the observed RT effects.

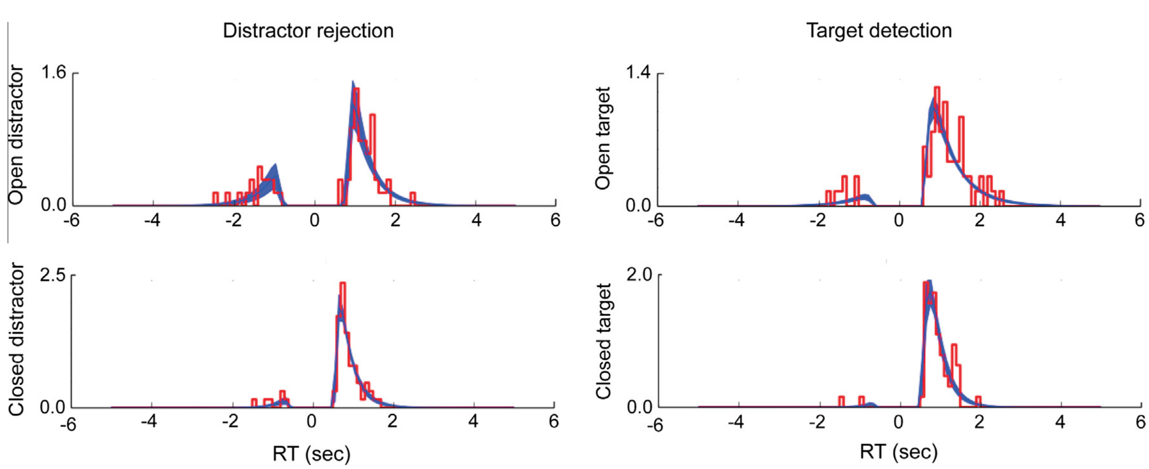


**Fig. 9.** Examples of the posterior predictive distribution as extracted from the optimal HDDM (blue lines), and respective empirical normalized RT distributions from one representative participant in Experiment 2 (red lines). Each panel depicts the distributions for separate (target-present) conditions in the experiment. Errors have been mirrored along the *x*-axis to display correct and incorrect RT distributions in one plot (positive and negative values, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

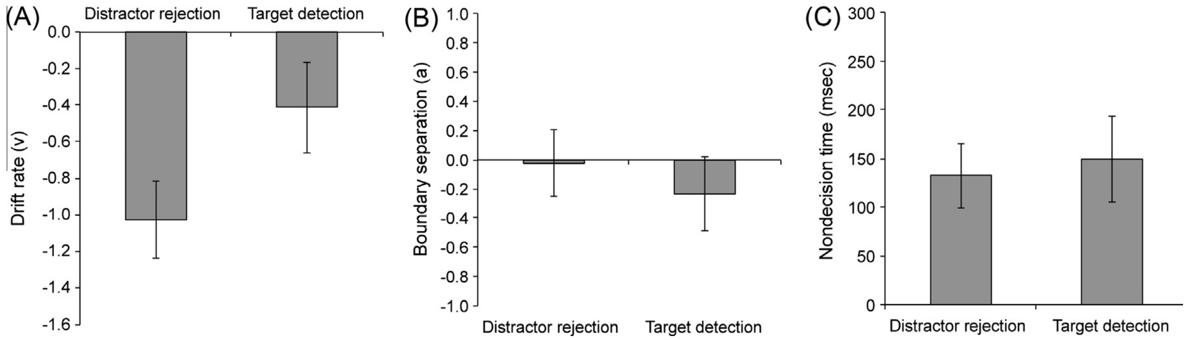


**Fig. 10.** Closure effect (open configurations minus closed configurations) for the mean hierarchical drift–diffusion parameters (A: drift rate; B: boundary separation; C: nondecision time) in Experiment 2 for the distractor rejection and the target detection task (comparing the respective target-present trials). The error bars represent ±1 standard error of the mean.

Differences in closure between targets and distractors were not reflected in nondecision times (Fig. 10C) or in the boundary separation parameter (Fig. 10B). Closed configurations afforded overall more efficient sensory encoding than open configurations, but there was no difference when comparing target- and distractor-related processing. Moreover, decision thresholds were comparable across open and closed configurations. Only the drift rates (Fig. 10A) showed a differential effect between distractor- and target-related processing, thus mirroring the RT pattern (Fig. 8D). This means that evidence accumulation was faster to reach the decision boundary for closed distractors than for open distractors. The same pattern was also observed when comparing closed and open targets, but the respective differences were substantially smaller.

## 4. General discussion

The current study aimed at elucidating how visual grouping by closure in targets and distractors contributes to the emergence of an 'illusory Gestalt'. To this end, two experiments were conducted employing a visual search task that presented variants of Kanizsa figures (Kanizsa, 1955), either inducing a 'part' or a 'whole' configuration with variations in grouping by closure. In Experiment 1, we found a robust CSE, that is, overall faster responses (by 180 ms) to wholes as compared to parts. Moreover, configural superiority was modulated by closure: detection of open targets (among closed distractors) showed a larger CSE than detection of closed targets (among open distractors; mean CSEs of 285 and 108 ms, respectively), with results being comparable for target-present and -absent trials. A diffusion model analysis on these data indicated that the observed CSE emerged at the stage of evidence accumulation. That is, a difference between closed and open configurations was revealed in the drift rate parameter, with faster evidence accumulation for wholes relative to parts with open targets (closed distractors), as compared to closed targets (open distractors). This pattern shows that the CSE in Experiment 1 primarily derived from processes related to the extraction of information to reach a decision. This process of information accumulation in turn seems to be particularly related to the suppression of closed, that is, well-grouped (distractor) configurations.

Next, in Experiment 2, we further investigated the role of grouping by closure, now systematically varying closure in targets and distractors independently of each other (using displays with whole-configurations only). Our analyses were primarily devised to compare the effect of closure in both targets and distractors, with closure quantified by subtracting search RTs for closed from RTs for open configurations. The results revealed a more robust effect of closure in distractor configurations as compared to targets (382 and 149 ms, respectively). Moreover, the enhanced closure effect in distractors was again reflected in the speed of evidence accumulation (the drift rate parameter). This analysis indicates that participants were overall faster to accumulate evidence for closed as compared to open configurations, but this benefit of closure was particularly pronounced with closure of distractor configurations, as compared to a much smaller effect with closure in the target configuration.

Taken together, the current results significantly extend previous studies on the CSE (Pomerantz & Portillo, 2011; Pomerantz et al., 1977) by showing that detection of a target configuration is facilitated primarily by the successful inhibition of distractors, with a considerably smaller role for target-related processing. While configural target processing may modulate search performance (Conci et al., 2011), in fact, we found no evidence of a reliable contribution of the target configuration to the CSE in Experiment 1. This suggests that configural superiority is not related to the emergence of an integrated object that matches a target description, or 'template', held in visual short-term memory. Such target templates are thought to have a privileged status, top-down biasing visual coding processes toward target-defining features (Olivers, Peters, Houtkamp, & Roelfsema, 2011). However, the current experiment yielded little evidence that the template status of the target is enhanced by object closure. Rather, the effect of the grouped configuration was particularly related to the distractors, suggesting that grouping by closure permitted more efficient suppression of task-irrelevant distractor configurations. One reason for the stronger effect of closure in distractors than in the target might simply derive from the fact that, in the typical CSE paradigm, there are more (most often three) distractors as compared to only a single target. In fact, visual search experiments show that, as set size increases, grouped distractors usually bring about a strong modulation of search efficiency (Conci et al., 2007a, 2007b), suggesting that the benefit of grouping in distractors increases as the number of candidate target configurations becomes larger (see also below and Humphreys & Müller, 1993). In this view, the 'emergence' of a configural target thus appears to be a by-product of the efficient suppression of a grouped array of distractors.

The critical stage that determined the observed pattern of the CSE was related to processes of evidence accumulation (as evidenced by the modulation of the drift-rate parameter), with closure in distractors speeding the rate of evidence accumulation. We propose that the emergence of a configural target from its constituent parts derives from the inhibition of distractor configurations. From this perspective, changes in the drift rate parameter are attributable to attentional control settings engaged in the inhibition of task-irrelevant objects, which are especially sensitive to the 'objecthood' (brought about by grouping mechanisms) in distractor arrangements (Kimchi, Yeshurun, & Cohen-Savransky, 2007).

The CSE has primarily been explained in terms of the Theory of Basic Gestalts (Pomerantz & Portillo, 2011), assuming a major role of perceptual grouping for the extraction of basic 'Gestalts', or emergent features, and treating such completed objects as the building blocks for perceptual organization. At the core of the theory is the formation of a Gestalt in a given object configuration, which permits faster and more efficient search for emergent features (that arise from the combination of parts into wholes on the basis of grouping) as compared to the corresponding basic features (i.e., properties of the parts, such as line orientation or color). Classical models of visual search (Treisman & Gelade, 1980; Wolfe, 2007) can usually not account for the CSE, and adding a uniform, non-informative context to search items would not normally be expected to improve performance (but rather only increase processing load). Context is usually added to all search items, but the relative contribution of targets and distractors in the build-up of emergent features has, to the best of our knowledge, not been investigated. In this regard, the current experiments reveal a preferential contribution of Gestalt formation to the CSE in visual search, which arises foremost from the distractors and only to a lesser extent from target-related processing.

Recent evidence suggests that a dedicated brain region in the ventral visual pathway, the LOC, may be particularly related to the processing of configurations, that is, emergent features (Kubilius et al., 2011). The authors showed that LOC (versus V1) was better able to predict the processing of wholes, whereas area V1 (versus LOC) better predicted the processing of part configurations. This pattern, showing processing of parts and wholes in distinct areas of the visual processing hierarchy, supports the idea that Gestalts may emerge only at a relatively high level of visual processing (beyond V1). Note that LOC has also been implicated in the processing of objects in general (Grill-Spector, Kourtzi, & Kanwisher, 2001) and illusory figures in particular (e.g., Bakar, Liu, Conci, Elliott, & Ioannides, 2008), for various tasks. In the light of our findings, the differences in the neuronal responses in LOC, as revealed by Kubilius et al. (2011), would appear to reflect the processing (in particular: the suppression) of distractor wholes, rather than the emergence of a configural target, thus resulting in a behavioral CSE.

In line with studies on the CSE, emergent features in illusory figures have been reported to yield efficient visual search performance (Davis & Driver, 1994; Gurnsey et al., 1992). Allocation of attention in search for Kanizsa-type figures is promoted, in particular, by grouping based on closure, that is, rendering a complete-object representation of the whole figure (Conci et al., 2006, 2007a, 2007b) – where implied closure is implicated in extracting a crude 'salient region' that can effectively guide search (for converging behavioral and electrophysiological evidence, see Conci et al., 2006, 2011; Wiegand et al., 2015). In this regard, the current findings suggest that the CSE for illusory figures is primarily related to distractor inhibition rather than target facilitation. Consistent with the present findings, a recent event-related potential (ERP) study has shown that search for a target Kanizsa figure can integrate information about distractors to optimize target selection (Töllner, Conci, & Müller, 2015) – suggesting that some form of distractor template drives top-down (distractor) suppression, thus reducing the distractors' impact on selection. In this view, the (relatively) efficient detection of a search target would be facilitated by the template-based rejection of grouped distractors (Duncan & Humphreys, 1989; see also Humphreys & Müller, 1993, for a computational model of template-based inhibition of distractors). For instance, Humphreys and Müller's model assumes that items (distractors, the target) compete to activate their respective templates, and in this competitive process, similar items (i.e., distractors of which there are multiple instances in the display) have a competitive advantage, that is, their template unit tends to cross the threshold first – upon which the whole set of distractors are 'rejected'. This is an essential component of the model and it might well account for the importance of closure in distractors, if one assumes that closed objects have an advantage in activating the respective template. Thus, these results and theoretical models are in accordance with the present findings, which provided evidence for the inhibition of closed distractor configurations, rather than facilitation of the corresponding targets, being the driving force of the behavioral CSE.

Distractor inhibition may not only operate at the level of grouped, configural objects, but also at that of basic features. For instance, a target defined by a simple, salient feature discontinuity (e.g., line orientation) in a field of uniform distractors usually leads to 'pop-out' (e.g., Müller, Heller, & Ziegler, 1995). One idea is that pop-out is the result of low-level local 'iso-feature' suppression (e.g., Zhaoping & May, 2007), that is, inhibitory interactions among nearby detectors coding similar features, impeding the distractors' ability to compete for selection and making the odd-ball (unsuppressed) target pop out. Recent ERP evidence suggests that such feature-based attention operates primarily via inhibition of distractor features, rather than activation of target features, at early stages of processing (Moher, Lakshmanan, Egeth, & Ewen, 2014). In this view, efficient detection of a target defined by a feature discontinuity is mediated by the suppression of uniform distractors, with potentially comparable mechanisms as described here for more complex object configurations.

Besides having a bearing on configural object processing and the CSE, the current results may also be seen as constituting a "search asymmetry" (Treisman & Gormican, 1988; Treisman & Souther, 1985; see also Wolfe, 2001). In a typical search asymmetry experiment, one of two stimuli (e.g., the letters Q and O) serves as target and the other as distractors in one condition (e.g., search for the Q among O's), with the target and distractor roles reversed in the other condition (e.g., search for the O among Q's). For this example, it has been shown that it is easier to find a target Q among distractor O's than finding a target O among distractor Q's (Treisman & Souther, 1985). The typical explanation for such an asymmetry is that, in the easier search condition, a distinctive feature (e.g., the stroke of the letter Q) would enable efficient search, while it is more difficult to find a target that is defined by the absence of a distinctive feature (e.g., the target O can be differentiated from the Q's as not having a stroke). The results of Experiment 1 obeys a comparable logic: We find more efficient performance when searching for an open target among closed distractors than when searching for a closed target among open distractors. However, in contrast to standard search asymmetries, this difference in performance does not arise because of a distinctive feature in the target (e.g., an emergent object that arises from grouping by closure), but rather the asymmetry results from the distinctive feature in distractors.

However, there are alternative explanations of search asymmetries in terms of distractor complexity. For instance, Rauschenberger and Yantis (2006) proposed that, in the above example (i.e., more efficient search with a Q target and O distractors than with the reverse assignment), the search asymmetry is caused not by (the presence vs. absence of) a distinctive feature in the target, but rather because O-shaped distractors are less complex stimuli than Q-shaped distractors, modulating search efficiency exclusively via distractor suppression (which is more efficient with less complex stimuli). Our findings lend support to this interpretation: less complex distractors (i.e., closed configurations) afford less effortful search than more complex distractors (i.e., open configurations).

Taken together, the present study points to a more prominent role of illusory Gestalt processing in the inhibition of distractors than previously thought, with implications for paradigms that investigate the role of inhibition in attention and awareness. For instance, in the perception of figure and ground, the assignment of a region in terms of being part of the figure or of the background determines which of the two leads to the prevailing percept – namely, the figure, while the other perceptual interpretation (of the background) is inhibited (e.g., Driver, Baylis, & Rafal, 1992; Roelfsema, 2006; Wagemans, Elder, et al., 2012). Moreover, in studies of binocular rivalry, where two incompatible stimuli are presented to each eye simultaneously, one of them will usually be temporarily suppressed in visual awareness, so as to make the other one perceived. Such interocular competition (between rivaling percepts) is solved by means of mutual inhibition enabling a single, coherent percept to emerge at any given moment in time (Kim & Blake, 2005). Thus, the present findings add to the notion that inhibition plays a major role in visual perception, in particular as regards the temporal and spatial filtering of the incoming sensory signals (Moors, Wagemans, van Ee, & de-Wit, in press; Tong, Meng, & Blake, 2011).

## 5. Conclusion

The present study reveals a major role of distractor inhibition in driving the emergence of an illusory Gestalt in Kanizsa figures. Our results show that the CSE is more pronounced when an emergent feature (e.g., as defined by closure) characterizes the search distractors rather than the target. Behavioral and drift–diffusion model evidence indicates that, in visual search, the configural superiority effect engendered by illusory figures arises primarily at the stage of evidence accumulation, where decisions are less driven by the conspicuity of the target configurations, but rather by the more effective suppression of grouped distractor configurations.

## Acknowledgments

This work was supported by project Grants from the German Research Foundation (DFG; CO 1002/1-1 and FOR 2293/1), and from the "LMUexcellent" Junior Researcher Fund.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.02.007.

## References


Bakar, A. A., Liu, L., Conci, M., Elliott, M. A., & Ioannides, A. A. (2008). Visual field and task influence illusory figure responses. *Human Brain Mapping, 29*(11), 1313–1326. http://dx.doi.org/10.1002/hbm.20464.

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

Conci, M., Böbel, E., Matthias, E., Keller, I., Müller, H. J., & Finke, K. (2009). Preattentive surface and contour grouping in Kanizsa figures: Evidence from parietal extinction. *Neuropsychologia, 47*(3), 726–732. http://dx.doi.org/10.1016/j.neuropsychologia.2008.11.029.

Conci, M., Gramann, K., Müller, H. J., & Elliott, M. A. (2006). Electrophysiological correlates of similarity-based interference during detection of visual forms. *Journal of Cognitive Neuroscience, 18*(6), 880–888. http://dx.doi.org/10.1162/jocn.2006.18.6.880.

Conci, M., Müller, H. J., & Elliott, M. A. (2007a). The contrasting impact of global and local object attributes on Kanizsa figure detection. *Perception & Psychophysics, 69*(8), 1278–1294. http://dx.doi.org/10.3758/BF03192945.

Conci, M., Müller, H. J., & Elliott, M. A. (2007b). Closure of salient regions determines search for a collinear target configuration. *Perception & Psychophysics, 69*(1), 32–47. http://dx.doi.org/10.3758/BF03194451.

Conci, M., Töllner, T., Leszczynski, M., & Müller, H. J. (2011). The time-course of global and local attentional guidance in Kanizsa-figure detection. *Neuropsychologia, 49*(9), 2456–2464. http://dx.doi.org/10.1016/j.neuropsychologia.2011.04.023.

Davis, G., & Driver, J. (1994). Parallel detection of Kanizsa subjective figures in the human visual system. *Nature, 371*, 791–793. http://dx.doi.org/10.1038/371791a0.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Psychological Science, 6*(3), 274–290. http://dx.doi.org/10.1177/1745691611406920.

Donnelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 561–570. http://dx.doi.org/10.1037/0096-1523.17.2.561.

Driver, J., Baylis, G. C., & Rafal, R. D. (1992). Preserved figure-ground segregation and symmetry perception in visual neglect. *Nature, 360*(73–75). http://dx.doi.org/10.1038/360073a0.

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

Eidels, A., Townsend, J. T., & Pomerantz, J. R. (2008). Where similarity beats redundancy: The importance of context, higher order similarity, and response assignment. *Journal of Experimental Psychology: Human Perception & Performance, 34*(6), 1441–1463. http://dx.doi.org/10.1037/a0012857.

Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research, 33*(7), 981–991. http://dx.doi.org/10.1016/0042-6989(93)90080-G.

Elder, J., & Zucker, S. W. (1994). A measure of closure. *Vision Research, 34*(24), 3361–3369. http://dx.doi.org/10.1016/0042-6989(94)90070-1.

Green, M. (1992). Visual search: Detection, identification, and localization. *Perception, 21*, 765–777. http://dx.doi.org/10.1068/p210765.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*(10–11), 1409–1422. http://dx.doi.org/10.1016/S0042-6989(01)00073-6.

Gurnsey, R., Humphrey, G. K., & Kapitan, P. (1992). Parallel discrimination of subjective contours defined by offset gratings. *Perception & Psychophysics, 52*(3), 263–276. http://dx.doi.org/10.3758/BF03209144.

Hout, M. C., Godwin, H. J., Fitzsimmons, G., Robbins, A., Menneer, T., & Goldinger, S. D. (2016). Using multidimensional scaling to quantify similarity in visual search and beyond. *Attention, Perception, & Psychophysics, 78*(1), 3–20. http://dx.doi.org/10.3758/s13414-015-1010-6.

Humphreys, G. W., & Müller, H. J. (1993). SEarch via Recursive Rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25*(1), 43–110. http://dx.doi.org/10.1006/cogp.1993.1002.

James, W. (1890). *The Principles of Psychology* (Vol. II). New York: H. Holt and Company.

Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea [Quasi-perceptual margins in homogeneously stimulated fields]. *Rivista di Psicologia, 49*(1), 7–30.

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible 'invisible'. *Trends in Cognitive Sciences, 9*(8), 381–388. http://dx.doi.org/10.1016/j.tics.2005.06.012.

Kimchi, R., Yeshurun, Y., & Cohen-Savransky, A. (2007). Automatic, stimulus-driven attentional capture by objecthood. *Psychonomic Bulletin & Review, 14*(1), 166–172. http://dx.doi.org/10.3758/BF03194045.

Koffka, K. (1935). *Principles of Gestalt psychology*. London, England: Lund Humphries.

Kogo, N., Strecha, C., Van Gool, L., & Wagemans, J. (2010). Surface construction by a 2-D differentiation–integration process: A neurocomputational model for perceived border ownership, depth, and lightness in Kanizsa figures. *Psychological Review, 117*(2), 406–439. http://dx.doi.org/10.1037/a0019076.

Kogo, N., & Wagemans, J. (2013). The emergent property of border-ownership and the perception of illusory surfaces in a dynamic hierarchical system. *Cognitive Neuroscience, 4*(1), 54–61. http://dx.doi.org/10.1080/17588928.2012.754750.

Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2011). Emergence of perceptual Gestalts in the human visual cortex: The case of the configural-superiority effect. *Psychological Science, 22*(10), 1296–1303. http://dx.doi.org/10.1177/0956797611417000.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., Wagenmakers, E. -J. (2015). JASP (Version 0.7). [*Computer software*].

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31*(3), 355–395. http://dx.doi.org/10.1016/0004-3702(87)90070-1.

Moher, J., Lakshmanan, B. M., Egeth, H. E., & Ewen, J. B. (2014). Inhibition drives early feature-based attention. *Psychological Science, 25*(2), 315–324. http://dx.doi.org/10.1177/0956797613511257.

Moors, P., Wagemans, J., van Ee, R., & de-Wit, L. (in press). No evidence for surface organization in Kanizsa configurations during continuous flash suppression. *Attention, Perception, & Psychophysics*.

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*(1), 1–17. http://dx.doi.org/10.3758/BF03211845.

Murray, M. M., & Herrmann, C. S. (2013). Illusory contours: A window onto the neurophysiology of constructing perception. *Trends in Cognitive Sciences, 17*(9), 471–481. http://dx.doi.org/10.1016/j.tics.2013.07.004.

Olivers, C. N., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends* *in Cognitive Sciences, 15*(7), 327–334. New York: Oxford University Press. http://dx.doi.org/10.1016/j.tics.2011.05.004.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442. http://dx.doi.org/10.1163/156856897X00366.

Pomerantz, J. R. (2003). Wholes, holes, and basic features in vision. *Trends in Cognitive Sciences, 7*(11), 471–473. http://dx.doi.org/10.1016/j.tics.2003.09.007.

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: Toward a theory of basic Gestalts. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1331–1349. http://dx.doi.org/10.1037/a0024330.

Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 635–649. http://dx.doi.org/10.1037/0096-1523.15.4.635.

Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance, 3*(3), 422–435. http://dx.doi.org/10.1037/0096-1523.3.3.422.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922. http://dx.doi.org/10.1162/neco.2008.12-06-420.

Rauschenberger, R., & Yantis, S. (2006). Perceptual encoding efficiency in visual search. *Journal of Experimental Psychology: General, 135*(1), 116–131. http://dx.doi.org/10.1037/0096-3445.135.1.116.

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29*, 203–227. http://dx.doi.org/10.1146/annurev.neuro.29.051605.112939.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B, 64*(4), 583–639.

Stanley, D. A., & Rubin, N. (2005). Rapid detection of salient regions: Evidence from apparent motion. *Journal of Vision, 5*, 690–701. http://dx.doi.org/10.1167/5.9.4.

Töllner, T., Conci, M., & Müller, H. J. (2015). Predictive distractor context facilitates attentional selection of high, but not intermediate and low, salience targets. *Human Brain Mapping, 36*(3), 935–944. http://dx.doi.org/10.1002/hbm.22677.

Tong, F., Meng, M., & Blake, R. (2011). Neural bases of binocular rivalry. *Trends in Cognitive Sciences, 10*(11), 502–511. http://dx.doi.org/10.1016/j.tics.2006.09.003.

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*(1), 15–48.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114*(3), 285–310.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin, 138*(6), 1172–1217. http://dx.doi.org/10.1037/a0029333.

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin, 138*(6), 1218–1252. http://dx.doi.org/10.1037/a0029334.

Wertheimer, M. (1912). Experimentelle studien über das sehen von bewegung [Experimental studies on the seeing of motion]. *Zeitschrift für Psychologie, 61*, 161–265.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift–diffusion model in Python. *Frontiers in Neuroinformatics, 7*. http://dx.doi.org/10.3389/fninf.2013.00014.

Wiegand, I., Finke, K., Töllner, T., Starman, K., Müller, H. J., & Conci, M. (2015). Age-related decline in global form suppression. *Biological Psychology, 112*, 116–124. http://dx.doi.org/10.1016/j.biopsycho.2015.10.006.

Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics, 63*(3), 381–389.

Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York: Oxford University Press.

Zhaoping, L., & May, K. A. (2007). Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *PLoS Computational Biology, 3*(4), e62. http://dx.doi.org/10.1371/journal.pcbi.0030062.

**Supplement**

Nie, Q.-Y., Maurer, M, Müller, H. J., & Conci, M. (2016) Inhibition drives configural superiority of illusory Gestalt: Combined behavioral and drift-diffusion model evidence, Cognition, 150, 150-162.

**Experiment 2 – Additional analyses on the mean RT and accuracy data**

The results of Experiment 2 in the main manuscript present the "closure effect" in the [target-present] RT and accuracy data by subtracting the averages of closed configurations from the corresponding averages of the open configurations. To complement these results, this supplement presents the analyses of both the target-present and – absent conditions without applying a subtraction procedure.

*Distractor rejection task*

Response accuracy. The mean percentage of correct responses from the distractor rejection task was calculated for each observer and variable combination. A 2x2 repeated-measures ANOVA on the percentage of correct responses, with the factors target presence (present, absent) and distractor closure (closed, open), revealed only the main effect of distractor closure (closed vs. open: 96% vs. 89%, $F(1,13) = 9.94$, $p = .008$, $\eta^2 = .43$, $BF_{10} = 177.7$) to be significant (Figure S1A). Neither the main effect of target presence nor the interaction between target presence and distractor closure was significant (all $ps > .35$, $\eta^2 s < .07$, $BFs < 0.38$).

Reaction times. An identical analysis was performed on mean RTs in the distractor rejection task (Figure S1A). The analysis again revealed only the main effect of distractor closure (closed vs. open: 744 vs. 1098 ms, $F(1,13) = 56.2$, $p < .001$, $\eta^2 = .8$, $BF_{10} = 1.21e+10$) to be significant. The main effect of target presence and the interaction between target presence and distractor closure were also not significant (all ps > .12, $\eta^2$s < .18, BFs < 0.38), mirroring the pattern in the accuracy data.

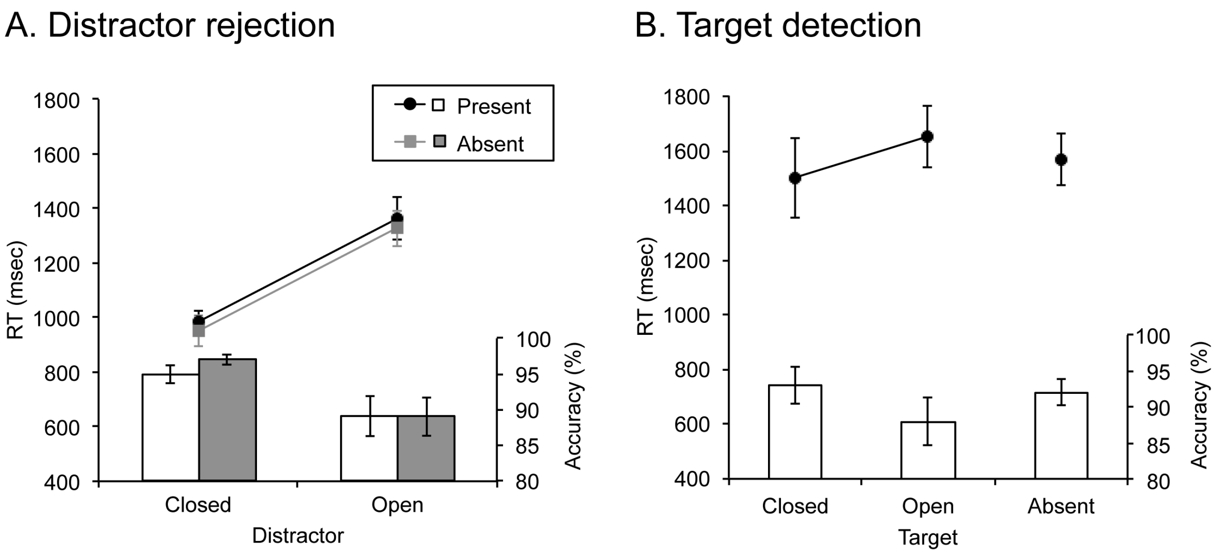


Figure S1. Behavioral results from Experiment 2. (A) Mean RTs (lines) and accuracies (bars) presented as a function of distractor closure for target-present and -absent conditions in the distractor rejection task. (B) Mean RTs (lines) and accuracies presented as a function of target closure (closed, open, absent) in the target detection task.

*Target detection task*

Response accuracy. Mean percentages of correct responses from the target detection task were calculated for each observer and condition. An one-way ANOVA on the percentage of correct responses, with the factor target closure (closed, open, absent) revealed no significant effect ($F(2,26) = 1.48$, $p = .25$, $\eta^2 = .43$, $BF_{10} = 0.57$; Figure S1B), and all post-hoc pairwise comparisons also showed no significant differences (all ps > .15, ds < .41, BFs < 0.7).

Reaction times. Mean RTs from the target detection task were analyzed similar to the above analysis on accuracies (Figure S1B). The analysis revealed no significant effect ($F(2,26) < 1$, $p = .39$, $\eta^2 = .07$, $BF_{10} = 0.34$). Notably, post-hoc paired t-tests nevertheless showed a significant difference between closed and open targets (closed vs. open: 1500 vs.

1649 ms, $t(13) = -2.33$, $p = .036$, $d = -0.62$, $BF_{10} = 2.01$), but no further significant differences ($ps > .41$, $ds < .23$, $BFs < 0.37$).

In sum, the pattern described here is essentially comparable to the outcomes presented in the main manuscript. Closed distractor configurations led to fewer errors and to faster responses than corresponding open distractor configurations. A comparable, but somewhat less reliable benefit was also revealed in the mean RTs for closed (as compared to open) target configurations.