Chen, S. and Nie, Q.-Y. and Muller, Hermann and Conci, M. (2019) Kanizsa-figure object completion gates selection in the attentional blink. Quarterly Journal of Experimental Psychology , ISSN 1747-0218. (In Press)

Kanizsa-figure object completion gates selection in the attentional blink

Siyi Chen, Qi-Yang Nie, Hermann J. Müller & Markus Conci

Ludwig-Maximilians-Universität München, Munich, Germany

Word count:9006 (main text), Abstract: 193

Correspondence:
Siyi Chen
Allgemeine und Experimentelle Psychologie
Department Psychologie
Ludwig-Maximilians-Universität
Leopoldstr. 13
D-80802 München
Germany
Email: Siyi.Chen@psy.lmu.de

Abstract

Previous work has demonstrated that perceptual grouping modulates the selectivity of attention across space. By contrast, how grouping influences the allocation of attention over time is much less clear. The current study investigated this issue, using an attentional blink (AB) paradigm to test how grouping influences the initial selection and the subsequent short-term memory consolidation of a target. On a given trial, two red Kanizsa-type targets (T1, T2) with varying grouping strength were embedded in a rapid serial visual presentation stream of irrelevant distractors. Our results showed the typical AB finding: impaired identification of T2 when presented close in time following T1. Moreover, the AB was modulated by the T2 grouping – independently of the T1 structure – with stronger grouping leading to a decreased AB and overall higher performance. Conversely, a reversed pattern, namely an increased AB with increasing grouping strength was observed when the Kanizsa figure was not task-relevant. Together, these findings suggest that the grouping benefit emerges at early perceptual stages, automatically drawing attentional resources, thereby leading to either sustained benefits or transient costs – depending on the task-relevance of the grouped object. This indicates that grouping modulates processing of objects in time.

Keywords: attentional blink, rapid serial visual presentation, perceptual grouping, Kanizsa figure.

Introduction

The organization of fragments and parts into coherent wholes is a central problem for visual perception. For instance, Kanizsa subjective figures (Figure 1A, complete; Kanizsa, 1955) give rise to a well-known visual illusion: the percept of an object with sharp contours and a brighter-than-background surface even though there is no actual luminance discontinuity in the physical stimulus. Kanizsa figures thus illustrate that the visual system can bind together separate parts (such as the "pacman" inducers in a Kanizsa figure) to produce a vivid impression of an integrated and coherent object. In this particular case, the association of distinct elements into a coherent whole has been shown to be governed by a set of Gestalt principles, such as collinearity and closure (Koffka, 1935; Wertheimer, 1923; see Brooks, 2015, for a recent review).

Visual search studies have consistently shown that component parts may be grouped prior to the engagement of attention (e.g., Moore & Egeth, 1997; Rensink & Enns, 1995). The critical measure in a visual search task is usually the time required to detect a particular target among a variable number of distractors. If the target is distinguished by a property that can be efficiently coded in parallel across the visual field, then it should "pop out", that is: search performance should not be affected by the number of distractors in the display. For instance, the search time for a target Kanizsa figure (Figure 1A, complete) is little affected by the number of distractor configurations (Figure 1A, ungrouped) that are composed of the same pacmen but do not induce an illusory figure (Conci, Müller, & Elliott, 2007a, 2009; Davis & Driver, 1994; Senkowski, Röttger, Grimm, Foxe, & Herrmann, 2005). Moreover, search for a Kanizsa target figure is far more efficient than search for a comparable "ungrouped" target configuration that does not render an illusory object, even though in both variants of the search task, the same distractor configurations were used, which were equally similar to both types of target (Conci et al., 2007a; Conci, Töllner, Leszczynski, & Müller, 2011). Together, these findings suggest that efficient search for Kanizsa figures is guided by grouping principles (i.e., collinearity and closure) that operate at early stages of visual processing, that is, prior to the engagement of attention (Conci, Müller, & Elliott, 2007b; Nie, Maurer, Müller, & Conci, 2016).

Integrated object configurations such as the Kanizsa figure have also been shown to automatically capture spatial attention. For example, search for a target disk in an array of randomly oriented (pacmen) distractor disks is substantially slowed when an illusory square is present (vs. absent) in the display (Rauschenberger & Yantis, 2001). Other experiments used search arrays containing a Kanizsa figure as a non-informative spatial 'cue' for a target that required a speeded choice reaction. Faster responses were obtained for a target presented inside, as compared to outside, the Kanizsa figure cue (Senkowski et al., 2005). Findings such as these suggest that a single integrated, illusory figure provides salient information, summoning an attentional orienting response to the region delineated by the grouped object (see also Marini & Marzi, 2016; Wiegand, Finke, Töllner, Starman, Müller, & Conci, 2015).

Whereas much of the previous work has elucidated how perceptual grouping modulates the allocation of selective attention across space, we know as yet little about how attentional selection is influenced by perceptual grouping over time. Temporal modulation of attention is frequently studied using the "attentional blink" (AB) paradigm, in which observers are asked to detect two targets presented successively within a rapid serial visual presentation (RSVP) stream of nontarget items (e.g., letters, words, or symbols) at a single location. While detection of the first target (T1) usually reveals a relatively high level of performance, detection of the second target (T2) is impaired if the temporal gap between the two targets is less than ~500 ms, while improving again at longer lags (Broadbent & Broadbent, 1987; Chun & Potter, 1995; Raymond, Shapiro, & Arnell, 1992). This transient drop in performance, which is referred to as the AB, has been assumed to reflect the temporal profile of attention.

The dual-target RSVP task can be thought of as a time-based analog of a visual search task (Vogel, Luck, & Shapiro, 1998). However, their underlying attentional mechanisms are not necessarily identical. Whereas processes of spatial attentional selection (e.g., in visual search) start to operate at an early, pre-attentive stage of processing, before stimulus identification is complete (see Luck, 1998, for review), the AB potentially reflects a post-perceptual attentional mechanism that marks the transition between perceptual stimulus analysis and the subsequent storage of selected items in a capacity-limited working memory buffer (Vogel et al., 1998). For instance, a prominent two-stage model to account for the AB (Chun & Potter, 1995) assumes that stage 1 involves perceptual coding of all stimuli in the RSVP stream; however, due to interference arising from the sequential mode of stimulus presentation, the encoded items decay rapidly over time, because each item is displaced by the one presented subsequently in the RSVP stream (see also Woodman & Luck, 2003; Moore & Lleras, 2005). To prevent or minimize interference, attentional resources are required to consolidate the “fragile” stimulus representations from stage 1 into a more stable and long-lasting format during stage-2 processing, that is, the consolidation of a selected number of items into working memory (see also Jolicœur & Dell’Acqua, 1998; Potter, Staub, & O’Connor, 2002; Shapiro, Raymond, & Arnell, 1997). Within this framework, an AB is thought to result from a failure of T2 to achieve stage-2 processing, because the capacity-limited consolidation mechanism is preoccupied with the processing of the preceding T1 stimulus (Vogel et al., 1998; Shapiro, Raymond, & Arnell, 1994).

Here, we investigated whether and how grouping structure in targets influences the profile of temporal attention. Although time-based selection operates only after initial visual processing, perceptual factors may nevertheless influence the AB (see e.g., Chen, Müller, & Conci, 2016, for effects of grouping on working memory). Previous studies, in fact, have shown that the AB is reduced when the perceptual salience of T2 is increased, for example, by increasing its featural and spatial dissimilarity to the distractors (Raymond, Shapiro, & Arnell, 1995) or by presenting highly arousing words (Anderson, 2005; Keil & Ihssen, 2004; Keil, Ihssen, & Heim, 2006) or familiar and emotional faces (Jackson & Raymond, 2006; Stein, Zwickel, Ritter, Kitzmantel, & Schneider, 2009) as T2s. To explain these findings, it has been suggested that salient stimuli are less susceptible to the AB because they generate a high level of (perceptual) activity that takes more time to decay, thus bridging the temporal gap during which resources are unavailable for encoding items into working memory (Anderson, 2005). On this background, we hypothesized that grouping in Kanizsa figures would lead to the formation of a salient object (Davis & Driver, 1994; Senkowski, et al., 2005; Conci, et al., 2007a; Rauschenberger & Yantis, 2001) that, in turn, would be relatively resistant against decay and more efficiently consolidated in spite of the limited capacity available, thus attenuating the AB.

To test this prediction, the present study investigated how perceptual grouping influences the AB using Kanizsa figures and comparable "ungrouped" control figures as targets. For instance, Experiment 1 implemented an RSVP stream of object configurations presenting circular placeholders in various colors. Observers were required to identify two uniquely colored (namely, red) target configurations. As illustrated in Figure 1A, the strength of grouping in the T2 configuration was gradually varied, ranging from a complete grouping (a Kanizsa star shape; Figure 1A left) through a partially grouped (three of six inducers form a Kanizsa triangle; Figure 1A middle) to an "ungrouped" configuration (no closure, all inducers point outwards; Figure 1A right) – thus systematically varying closure in the Kanizsa-type configurations. Note that the various pacman configurations changed in terms of the strength of grouping they engendered, however without changing the low-level properties of the image. The crucial question concerned whether the accuracy of identifying the T2 target configuration would vary as a function of its grouping strength. That is, by systematically varying the T1–T2 lag, we examined whether the grouping structure of T2 would modulate the AB effect.

Experiment 1

Experiment 1 was performed to investigate how T2 grouping strength influences the AB. On a given trial, distractor arrangements of six colored disks (all disks of the same color, but not red) were presented in rapid succession. Within this stream, two arrangements were presented in red and these were defined as the target configurations. Targets were presented with small segments removed from each disk, which, by appropriately rotating the cutout segments, would create the impression of a Kanizsa figure. T1 was always defined as a grouping (of cutout disks) that would not lead to the emergence of an illusory figure, and T2 was either a complete (Kanizsa star shape), or a partially grouped (Kanizsa triangle shape), or an ungrouped configuration (see Figure 1A). This manipulation permitted us to examine whether a systematic variation of the grouping strength in T2 would influence the AB. We predicted that the AB effect would be dependent on the grouping strength of the T2 configuration, with a reduced AB for more strongly grouped T2 objects.

Method

Participants. Fifteen right-handed volunteers (7 male; mean age: 24.67 ± 2.26 years) with normal or corrected-to-normal visual acuity and (self-reported) normal color vision participated in the experiment for payment of €8.00 per hour. The experimental procedure was approved by the ethics committee of the Department of Psychology, Ludwig-Maximilians-Universität München, and all participants provided written informed consent. Sample size was determined on the basis of previous, comparable studies (e.g., Stein et al., 2009), aiming for 85% power to detect a medium effect size (within-participants; f=0.25; Cohen, 1988) given an alpha level of .05.

Apparatus. The experiment was conducted on a Windows computer using Matlab routines and Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli were presented in different colors against a gray (RGB: 125, 125, 125; 51.7 cd/m^2) background in the center of a 17-inch computer monitor (1024×768 pixel screen resolution, 85-Hz refresh rate). The experiment was conducted in a sound-attenuated room that was dimly lit with indirect incandescent lighting.

Stimuli. Each trial consisted of a series of configurations that were presented in different colors. Each configuration was composed of six colored disks (each subtending a viewing angle of 3.3°) arranged around a circle (with a radius of 5.2°, at a viewing distance of 50 cm). Distractor configurations were composed of six complete disks presented in four different colors – green (RGB: 75, 184, 72), purple (RGB: 137, 41, 143), yellow (RGB: 243, 236, 27), or blue (RGB: 22, 148, 210) – which were selected at random, with the only restriction that two consecutive configurations never shared the same color. Two target configurations composed of six pacmen inducers (i.e., disks with quarter segments removed) were presented in red color (RGB: 236, 30, 39). The first target (T1) was presented with all pacmen inducers oriented either up- or downwards (see Figure 1A). The second target (T2) either presented a complete Kanizsa figure (a star shape), a partially grouped Kanizsa triangle (with the triangle presented in up- or downward orientation), or an ungrouped object (with all pacmen inducers rotated outwards by 180°), thus gradually varying the grouping strength of T2 by means of a decrease in object closure (see Figure 1A).

Procedure and Design. As depicted in Figure 1B, each trial started with the presentation of a central fixation cross for 500 ms at the screen center, followed by the RSVP stream. Each configuration was presented for 100 ms, followed by a short blank interval of 20 ms, resulting in a stimulus onset asynchrony (SOA) of 120 ms. On a given trial, the first target (T1) was randomly allocated to one of three temporal serial positions, from 2 to 4, within the stream of 10 configurations. The second target (T2) was presented at one of four different temporal lags (120, 240, 360, or 600 ms, corresponding to serial lag positions 1, 2, 3, or 5) subsequent to T1. RSVP stream distractors continued to be presented during the lag and after T2. Following the presentation of the RSVP stream, a blank screen appeared until a response was issued. Participants were instructed to detect the two red targets. With regard to T1, participants were asked to identify the pointing direction of the T1 pacmen with a right (upward) or left (downward) mouse button press, using (the index or the middle finger of) the right hand. With regard to T2, the task was to report the number of triangles that were presented within a given target configuration, that is: “0” for the ungrouped object, “1” for the partially grouped object, and “2” for the complete object grouping. Participants responded with a left-hand button press via keyboard, pressing the left-, down-, or rightward arrow key for “0”, “1”, or “2” triangles, respectively. Feedback was provided at the end of each trial by displaying a white and/or a red “—” sign for 500 ms on the screen if an error had occurred for the first and/or the second target, respectively. Trials were separated from one another by an interval of 1000 ms. Observers were instructed to respond as accurately as possible, with particular emphasis on T1 accuracy in order to maximize the number of trials available for the analysis of T2 accuracy. Every participant completed 20 blocks of 24 trials each, following one practice block of 24 trials (giving a total of 504 trials). After each block, participants took a short break; they then proceeded to the next block by pressing the ‘space’ bar following a message on the screen.

In summary, the experiment systematically varied two factors: T2-target type (complete, partially grouped, and ungrouped object), and T1–T2 lag (Lag 1, 2, 3, or 5), with all possible factorial combinations presented in random order.

Figure 1 about here

Results

T2 accuracy. Estimates of T2 accuracy were based solely on trials on which T1 had been identified correctly (as it is hard to determine the effect on the processing of T2 when the cause of the erroneous response to T1 is not known). Figure 2 presents T2 accuracy as a function of lag, separately for the different target conditions. A two-way repeated-measures analysis of variance (ANOVA) of correct T2 responses, with the factors T2-target type (complete, partially grouped, ungrouped) and lag (1, 2, 3, 5), revealed both main effects to be significant: T2-target type, $F(2, 28) = 6.67$, $p = .004$, $\eta_p^2 = .32$, 90% confidence interval, or CI [.07, .48]; lag, $F(3, 42) = 24.12$, $p < .0001$, $\eta_p^2 = .63$, 90% CI [.45, .71]. For the post-hoc comparisons, given that such repeated testing increases the chance of a significant effect, a Bonferroni correction was applied (Neter & Wasserman, 1974). There was a graded effect of target type, with the highest accuracy for complete configurations (86%), followed by partially grouped (81%) and ungrouped (74%) configurations (complete vs. ungrouped: $p = .001$; partially grouped vs. ungrouped: $p = .39$; partially grouped vs. complete: $p = .31$). In addition, there was a monotonic increase of performance from lag 1 onwards (76%, 77%, 83%, and 87% for lags 1, 2, 3, and 5, respectively; *p*s < .029).

Furthermore, the T2-target type × lag interaction was significant, $F(6, 84) = 2.68$, $p = .02$, $\eta_p^2 = .16$, 90% CI [.01, .23]. To decompose this interaction, the AB amplitude was computed (see also Anderson, 2005), which is defined as the maximum difference in performance across lags, that is, contrasting the (early) lag(s) with the lowest accuracy with (later) lag(s) that resulted in the highest level of accuracy (in Experiment 1, the largest difference was revealed between lag 1 and lag 5). Comparisons of the AB amplitude across target type conditions revealed the difference in amplitude to be largest for ungrouped (15%), intermediate for partially grouped (11%), and smallest for complete configurations (6%), $F(2, 28) = 3.47$, $p = .045$, $\eta_p^2 = .20$, 90% CI [.00, .36].

T1 accuracy. The mean percentage of correct responses for T1 was 90%. A two-way repeated measures ANOVA with the factors T2-target type (complete, partially grouped, ungrouped) and lag (1, 2, 3, 5) revealed only a lag effect: $F(3, 42) = 35.49$, $p < .0001$, $\eta_p^2 = .72$, 90% CI [.56, .78]. T1 performance exhibited a drop at lag 1: 80%, 92%, 93%, and 94% for lags 1, 2, 3 and 5, respectively (*p*s < .0001); that is, the short lag between T1 and T2 (also) impacted the accuracy of reporting T1. Importantly, however, no (main or interaction) effect involving T2-target type was revealed (*F*s < 1.5, *ps* > .25).

Figure 2 about here

Discussion

The pattern of results in Experiment 1 clearly demonstrates that T2-targets are the less susceptible to the AB the higher their grouping strength: the AB amplitude was smallest for the complete, intermediate for partially grouped, and largest for ungrouped T2 configurations. In addition, the overall T2 accuracy also depended on the grouping strength, with higher performance for the more strongly grouped objects. Importantly, there was no influence of the T2 grouping type on T1, that is, the enhanced accuracy for complete and partially grouped T2s cannot be explained in terms of a trade-off between T2 and T1 accuracy. Our finding that grouping was associated with a diminished AB suggests that attention was more effectively allocated to grouped stimuli, in line with our initial prediction. This is consistent with findings from previous studies on the spatial allocation of attention, which have been taken as evidence for the preattentive coding of Kanizsa figures (e.g., Davis & Driver, 1994; Senkowski, et al., 2005; Conci, et al., 2007a; Rauschenberger & Yantis, 2001). This benefit of grouping manifested even though the complete and ungrouped objects consisted of identical physical stimulus components and were of equal object complexity (in terms of the descriptive criteria of Garner & Clement, 1963). As regards attentional guidance, the preattentive formation of a global object representation is beneficial even given minor variations in terms of the geometry and spacing of the local inducer elements that make up a given stimulus configuration (Chen, Glasauer, Müller, & Conci, 2018; Conci et al., 2007b).

However, there still remained a lag-dependent impairment for the grouped T2, which (although the decrement became smaller with increasing lag) would appear to be at variance with the view that the illusory shape is processed completely independently of attention (see also Joseph, Chun, & Nakayama, 1997). A potential explanation for selection of a grouped object being to some extent dependent on attention refers to the idea that featural and configural information are processed in somewhat different ‘channels’ (Awh et al., 2004) and that interference would arise to the degree to which T1 and T2 overlap in terms of the processing channels involved. In the present experiment, both T1 and T2 share information at the feature level (i.e., they consist of the comparable inducer elements) – which would engender a degree of feature-based interference, resulting in an AB. However, the grouped T2 object would additionally be processed via the configural channel. This would reduce the total interference as grouping strength increases, especially when a global object emerges.

Note that Experiment 1 revealed monotonic increases in T2 identification with longer T1–T2 lags, while many previous studies have reported an effect of “lag-1 sparing” in which performance is relatively unimpaired if T2 is presented directly after T1 (e.g., Chun & Potter, 1995; Raymond et al., 1992). A potential explanation for this sparing effect is that the visual system tends to process the two targets together (e.g., in a batch) as long as they appear in direct temporal succession (Chun & Potter, 1995). However, it has also been shown that lag-1 sparing occurs in particular when no attentional switch (e.g., across locations, tasks, or categories) is required between targets (Kawahara, Zuvic, Enns, & Di Lollo, 2003; Di Lollo, Kawahara, Ghorashi, & Enns, 2005; Visser, Bischof, & Di Lollo, 1999). Lag-1 sparing is in addition crucially dependent on the temporal separation between targets, with reliable sparing being evident predominantly with lags shorter than 100 ms (Olivers & Meeter, 2008; Potter, Staub, & O'Connor, 2002). The results from Experiment 1 failed to show spared lag-1 performance; rather, the AB was particularly pronounced at lag 1. This may have resulted from the task switch between two targets (from a local-object direction discrimination task to a global-shape "counting" task) and from the relatively long T1–T2 lag (120 ms; see also Conci & Müller, 2009).

Experiment 2

Experiment 1 showed that T2 grouping strength modulates the AB when T1 is an ungrouped configuration that requires the identification of the (individual) pacman's pointing direction. In Experiment 2, we investigated whether grouping in T1 might also influence T2 processing. This was motivated by findings that the AB may actually be increased following a salient T1 (Martens & Wyble, 2010; i.e., the converse of the reduction of the AB by a salient T2). This has been attributed to the increased salience of T1 engendering a longer dwell of attention (on the T1) and thus reducing the capacity available for the processing of T2 (Stein et al., 2009; Huang, Baddeley, & Young, 2008). In Experiment 2, we therefore increased the strength of the T1 grouping by presenting a partially grouped Kanizsa triangle in order to examine whether the selection of a grouped T1 would impede the consolidation of complete, partially grouped, and ungrouped T2 configurations. Grouped Kanizsa figures have previously been shown to capture attention (see Introduction). Accordingly, we expected a salient T1 Kanizsa figure to lead to an overall increase of the AB. Moreover, when assuming that the (grouped) T1 stimulus is processed via separate, featural and configural "channels" (Awh et al., 2004), grouping in T2 should be associated with a reduced benefit when – that is, there should be a rather strong AB for all types of stimulus. By contrast, a single processing "channel" account (as in Chun & Potter, 1995) would predict a strong effect of grouping in T2 (as in Experiment 1), because the grouped T2 would nevertheless be more likely to escape the AB than an ungrouped T2.

Recall that the AB modulation by means of the T2 grouping strength in Experiment 1 was maximal at early temporal lags, but a substantial difference between configuration types nevertheless remained until later lags. For instance, the complete T2-target gave rise to a significantly higher accuracy than the ungrouped T2 across all lags (significant main effect of T2-target type), and this difference persisted even until lag 5, that is, 600 ms after the presentation of T1, $t(14)=2.56$, $p = .023$, $d = .66$, 95% CI [.09, 1.21]. A potential explanation for this sustained difference between T2 groupings might be that the temporal interval between T2 and T1 was simply not long enough, even at lag 5; that is, selection of T2 some 600 ms after T1 might still be compromised due to the attentional demands of processing the preceding T1. However, an alternative explanation might be that the benefit of grouping at longer lags reflects an additional advantage that arises from post-selective processing (i.e., at stage 2). In this view, how efficiently a given target configuration is consolidated into short-term memory would vary for the various types of grouping. To address this issue, in Experiment 2, the temporal lags were extended (beyond lag 5) up to lags 6 and 7. More precisely, T2 was presented at one of four different temporal lags (120, 240, 720, or 840 ms, corresponding to serial lag positions 1, 2, 6, or 7), thus covering an extended time interval subsequent to T1.

Method

Experiment 2 was methodologically identical to Experiment 1, except that the T1 configuration was always a partial grouping that induced a Kanizsa triangle which pointed either up- or downwards (see Figure 3). The T1 task was roughly comparable to Experiment 1: it required observers to identify the pointing direction of the triangular T1 grouping (up- vs. downwards). With respect to T2, observers had again to determine the number of triangles (as in Experiment 1). In addition, compared to Experiment 1, the T1–T2 lags were extended. On a given trial, T1 was randomly allocated to one of three temporal serial positions, from 2 to 4, within a stream of now 12 configurations. T2 was then presented at one of four different temporal lags (120, 240, 720, or 840 ms, corresponding to serial lag positions 1, 2, 6, or 7) subsequent to T1. As in Experiment 1, RSVP stream distractors continued to be presented during the lag and after T2. A new group of fifteen right-handed volunteers (7 males; mean age: 23.00 ± 2.83 years) with normal or corrected-to-normal visual acuity participated in the experiment for payment of € 8.00 per hour. Each participant completed 24 practice plus 480 experimental trials (divided into 20 blocks).

Figure 3 about here

Results

T2 accuracy. Figure 4 presents the T2 accuracy (given a correct T1 response) as a function of lag, separately for the different target type conditions. A two-way repeated-measures ANOVA of correct T2 responses with the factors T2-target type (complete, partially grouped, ungrouped) and lag (1, 2, 6, 7) revealed both main effects to be significant: target type, $F(2, 28) = 14.12$, $p < .0001$, $\eta_p^2 = .50$, 90% CI [.24, .63] and lag, $F(3, 42) = 28.80$, $p < .0001$, $\eta_p^2 = .67$, 90% CI [.50, .74]. T2 accuracy was higher for complete (85%) than for partially grouped (73%; $p = .004$) and ungrouped (71%; $p < .0001$) configurations; there was no significant accuracy difference between partially grouped and ungrouped configurations ($p = 1$). Moreover, T2 accuracy increased with T1–T2 lag (67%, 69%, 84%, and 86% for lag 1, 2, 6, and 7, respectively), revealing a significant increase from lag 2 onwards (ps <.001), but no significant difference for the lag-1 vs. lag-2 comparison ($p = 1$). In addition, the T2 target type × lag interaction was significant, $F(6, 84) = 2.34$, $p = .039$, $\eta_p^2 = .14$, 90% CI [.00, .21], mainly due to a performance difference between the complete and ungrouped condition, $F(3, 42) = 6.88$, $p = .001$, $\eta_p^2 = .33$, 90% CI [.11, .45]: the AB amplitude (lags 1/2 vs. 7) was larger for ungrouped (22%) compared to complete configurations (13%), $t(14) = 3.01$, $p = .009$, $d = .78$, 95% CI [.19, 1.35]. The partially grouped configuration exhibited an intermediate AB amplitude (20%), but this did not differ from the ungrouped (p = .67) or complete (p = .29) configurations.

Figure 4 about here

T1-T2 pointing direction. In a subsequent analysis, we examined whether the (up-/downward) pointing direction of the partially grouped triangle in T1 influenced the detection performance for the (up-/downward pointing) T2 in partially grouped configurations. Figure 5B presents T2 accuracy as a function of lag, separately for the same and different orientations of the Kanizsa triangles. A two-way repeated-measures ANOVA of correct T2 responses with the factors T1–T2 direction (same, different) and lag (1, 2, 6, 7) revealed all main effects to be significant: T1–T2 direction, $F(1, 14) = 47.83, p < .0001, \eta_p^2 = .77$, 90% CI [.52, .85] and lag, $F(3, 42) = 14.05, p < .0001, \eta_p^2 = .50$, 90% CI [.28, .60]. T2 accuracy was higher for matching than for mismatching pointing directions (80% vs. 65%). T2 accuracy increased with T1–T2 lag, as described above. The interaction was also significant, $F(3, 42) = 3.08, p = .038, \eta_p^2 = .18$, 90% CI [.00, .30]: the accuracy difference between matching and mismatching pointing directions was reliable for the first three lags (ps < .003), but no longer reliable (i.e., reduced) at lag 7 (p = .07). Thus, in Experiment 2, the orientation similarity of the (Kanizsa) triangles modulated performance.

An analogous analysis was also performed for Experiment 1 (Figure 5A), comparing the same/different pointing direction of the T1 pacmen and the subsequent T2 triangle configuration. This analysis revealed only a significant main effect of lag, $F(3, 42) = 17.11, p < .0001, \eta_p^2 = .55$, 90% CI [.34, .64], illustrating the AB effect pattern already described above (for Experiment 1). The fact that there was no effect of the same/different pointing direction at any lag (all ts(14) < 1.35, ps > .20; see Figure 5A) means that, in contrast to Experiment 2, there was no influence of the local pacman direction in T1 on the detection of T2 triangles in Experiment 1.

Figure 5 about here

Cross-experiment comparison. To examine whether the change of the T1 target across Experiments 1 and 2 affected the AB and processing of the grouped/ungrouped T2 targets, we compared the AB amplitude between the two experiments in a mixed-design ANOVA with the within-subject factor T2-target type (complete, ungrouped) and the between-subject factor Experiment (1, 2). This analysis revealed a significant main effect of T2-target type, $F(1, 28) = 15.25$, $p < .001$, $\eta_p^2 = .35$, 90% CI [.12, .52], with an overall larger AB amplitude for ungrouped (19%) than for complete (10%) T2 configurations. There was also a marginally significant main effect of Experiment, $F(1, 28) = 3.12$, $p = .08$, $\eta_p^2 = .1$, 90% CI [.00, .28], reflecting a somewhat larger AB amplitude in Experiment 2 (18%) than in Experiment 1(11%). The interaction was not significant ($F < 1$, $p > .8$).

T1 accuracy. Accuracy of T1 identifications was again relatively high, with an average of 90% correct responses, comparable to T1 performance in Experiment 1, $t(28) = .33$, $p = .75$, $d = .12$, 95% CI [-.60, .84]. A two-way repeated-measures ANOVA with the factors T2-target type (complete, partially grouped, ungrouped) and lag (1, 2, 6, 7) only revealed a significant main effect of lag, $F(3, 42) = 13.60$, $p < .0001$, $\eta_p^2 = .49$, 90% CI [.27, .60]), with accuracy being reduced at lag 1 (86%, 90%, 93%, and 92% for lags 1, 2, 6 and 7; ps < .003), comparable to the finding in Experiment 1. There were no other significant effects (Fs < 1, ps > .35).

Discussion

The results of Experiment 2, in general, replicate those of Experiment 1, in that performance was overall reduced and the AB amplitude was larger for ungrouped relative to complete-object T2 configurations. Moreover, a comparison between Experiments 1 and 2 showed that increasing the strength of the T1 grouping translated into a somewhat increased AB overall. This pattern suggests that the effect of grouping on T2 detection is largely independent of the perceptual structure of the T1 stimuli, even though increasing the salience of T1 (in the present experiment: from "ungrouped" arrangements of pacmen to a coherent illusory triangle) leads to an increased difficulty in the processing of T2, because of a prolonged dwell of attention on T1. This outcome is hard to explain on the assumption of separate featural and configural processing channels (Awh et al., 2014), because the grouped T1 would have occupied both channels, thus reducing the impact of grouping in T2. Instead, the current results would appear to be more compatible with the assumption of a single channel (as, e.g., in Chun & Potter, 1995), according to which the salient T2 grouping would lead to a modulation of performance that is largely independent of the T1 structure.

Despite the lag × T2-target type interaction, there was still a significant difference between the completed and ungrouped T2 at both shorter lags (*p*s = .0001) and longer lags (*p*s < .001), which mirrors the result pattern of Experiment 1. For instance, even with a T1–T2 separation of 840 ms (at lag 7), performance for the ungrouped T2 configuration was still reduced relative to performance for T1 (mean difference: -7.04; $p < .04$). By contrast, performance for the complete T2 was roughly comparable (if not, in fact, being somewhat higher compared) to performance for T1 (mean difference: 2.96; $p = .06$). This suggests that the reduced performance for the ungrouped T2 does not solely reflect the temporal dynamics of attentional selection, that is, a sustained difficulty in selecting T2 while being engaged with T1. Rather, this constant difference across groupings might point to a difference in the efficiency with which the ungrouped vs. the complete T2 is retained at a post-selective stage in short-term memory until the execution of the response.

A second influence of T1 processing on T2 performance was revealed by the analysis of the same/different triangle pointing directions across the T1 and T2 (partially grouped) targets: accuracy was higher for T2 when the T2 triangle orientation matched that of T1, while accuracy was lower when they mismatched. No analogous effect was obtained in Experiment 1, in which the pacmen's local orientation and the global orientation of the triangle grouping could repeat across T1 and T2.

One might argue that responding to T1 in Experiment 2 would not necessarily require the completion of an up- or downward-pointing triangle, but that instead the response might solely be based on the up-/downward-pointing indentation of a single pacmen inducer, for example, the upper pacman in the T1 configuration (comparable to the task in Experiment 1). However, this seems rather unlikely given the different result patterns from the analysis of the (same vs. different) pointing directions across experiments. In Experiment 1, judging the orientation of ungrouped inducer elements (T1) did not influence the extraction of a grouped triangle (T2); in Experiment 2 by contrast, the extraction of a grouped T1 triangle substantially influenced the subsequent processing of the grouped T2 triangle. This indicates that performance was not simply modulated by some form of response priming between T1 and T2. Instead, observers did complete the presented shapes, and they did perform the task in line with the instructions provided.

The finding of a same-object benefit for identical T1 and T2 stimuli in Experiment 2 is also consistent with Raymond (2003; see also Conci & Müller, 2009). Our results mirror these previous findings and further show that repeated perceptual objects (Experiment 2), rather than repeated response-defining features (Experiment 1), lead to a reduction of the AB. Note that repeating the perceptual objects from T1 to T2 led to an attenuation but not to complete abolishment of the AB. This “residual” AB might have resulted from the change in task demands from T1 to T2 (see Visser, Bischof, & Di Lollo, 1999).

Finally, it should be noted that performance for the “different” (up-/downward pointing direction) condition was relatively low (65%), which may, to some extent at least, be attributable to a variant of “accidental” binding (Karabay & Akyürek, 2017). On this account, the presentation of two triangles pointing in opposite directions might yield the erroneous percept of a single Kanizsa star, integrating the sequential triangles into a unitary configuration. Such erroneous bindings would be particularly prominent at short temporal lags. To examine for this, we computed the frequency of participants reporting an integrated percept (i.e., a Kanizsa star) for the partially grouped target, given different T1 and T2 orientations. Indeed, erroneous Kanizsa star reports were rather frequent at lag 1 (36%), and declined at longer lags (21%, 13%, and 11% for lags 2, 6, and 7, respectively), revealing a linear trend: $F(1, 14) = 15.36$, $p = .002$, $\eta_p^2 = .52$, 90% CI [.17, .69]. This is consistent with observers tending to merge the two opposite triangles presented in succession into a single, coherent representation – consistent with the notion of “misbinding”. For the “same” condition, by contrast, the erroneous star reports were significantly reduced (compared to the "different" condition), F (1, 14) = 23.50, $p < .0001$, $\eta_p{}^2 = .63$, 90% CI [.29, .75], revealing overall comparable rates of erroneous star reports across lags (12%, 11%, 9%, and 8% for lags 1, 2, 6, and 7, respectively, ps > .28).

Experiment 3

In the experiments presented thus far, participants were not just passively exposed to variants of Kanizsa figures (with varying grouping strength), but they were required to actively classify these configurations, that is, to report the number of triangles presented in T2. Both experiments demonstrated a comparable pattern of results, namely a diminished AB and enhanced performance across all lags when T2 presented a complete (as opposed to an ungrouped) configuration. This pattern was obtained regardless of the type of object presented as T1, suggesting some automaticity in processing the grouped object. Experiment 3 was performed to further elucidate how the specific task to classify a given object configuration in T2 determined the grouping effect. To investigate this issue, in Experiment 3, the T2 task was changed such that the requirements were unrelated to the object configuration presented. This was achieved by adding a small arrow (an oriented ">"-sign) to the (complete, partially grouped, or ungrouped) T2 configuration, and the T2 task was to report the orientation of the unrelated arrow while the grouping itself was essentially task-irrelevant (see Figure 6). It should be noted that the (red) color of the pacman inducers still acted as a target cue, intended to ensure that observers processed the stimulus, but the (color) cue was completely independent of the grouping structure displayed in T2. We expected that if grouping does require top-down attention, then the change of the task requirements (in Experiment 3) should eliminate the above AB modulation of grouping (as attention does no longer need to be paid to the grouping, but only to the task-relevant arrow). By contrast, if grouping engenders automatic, early perceptual processing, then one would expect that the T2 accuracy would still be modulated by the (in Experiment 3) entirely task-irrelevant groupings.

Method

Experiment 3 was similar to Experiment 1, except that the (complete, partially grouped, or ungrouped) T2 configuration was now presented for 70 ms, after which a small arrow (0.5° × 0.5°) was added to the RSVP stream for another 30 ms (see Figure 6). The presentation duration of the arrow (target) was determined based on pilot tests, which showed that a relatively short presentation time is necessary to guarantee a reasonable variability of performance (i.e., well below ceiling). As in Experiment 1, the presentation of the stimuli was followed by a 20-ms blank interval, yielding a 120-ms SOA as in Experiments 1 and 2. The T2 task was to report the up/down/left/right pointing direction of the arrow, which was randomly presented at three possible locations within a given configuration (i.e., at top-left, top-right, or bottom locations; see Figure 6). Participants responded with a left-hand button press via keyboard, pressing the corresponding up-, down-, left-, or rightward-pointing arrow key, respectively. On a given trial, T1 was randomly allocated to one of three temporal serial positions, from 2 to 4, within a stream of 12 configurations. T2 was then presented at one of four different temporal lags (120, 240, 720, or 840 ms, corresponding to serial lag positions 1, 2, 6, or 7) subsequent to T1 (i.e., the lags were the same as in Experiment 2). RSVP stream distractors continued to be presented during the lag and after T2. The T1 target and task and the distractors remained the same as in Experiment 1. Fifteen naive, right-handed volunteers (7 males; mean age: 23.67 ± 2.66 years) with normal or corrected-to-normal visual acuity participated in the experiment for payment of 8.00 Euro per hour. Each participant completed 24 practice plus 480 experimental trials (divided into 20 blocks).

Figure 6 about here

Results

T2 accuracy. Figure 7 presents the T2 accuracy (given a correct T1 response) as a function of lag, separately for the different target type conditions. A two-way repeated-measures ANOVA of correct T2 responses, with the factors T2-target type (complete, partially grouped, ungrouped) and lag (1, 2, 6, 7), revealed both main effects to be significant: T2-target type, *F*(2, 28) = 5.40, $p = .01$, $\eta_p^2 = .28$, 90% CI [.04, .44], and lag, *F*(3, 42) = 12.00, $p < .0001$, $\eta_p^2 = .46$, 90% CI [.24, .57]. There was a graded effect of target type, with the highest accuracy for ungrouped configurations (96%), followed by partially grouped (95%) and complete (94%) configurations (complete vs. ungrouped: $p = .02$; partially grouped vs. ungrouped: $p = .40$; partially grouped vs. complete: $p = .29$). In addition, there was a monotonic increase in performance from lag 1 onwards (92%, 94%, 97%, and 98% for lags 1, 2, 6, and 7, respectively; *p*s < .017, except for comparable performance with lags 6 and 7, $p = .56$). The T2-target type × lag interaction was also significant, *F*(6, 84) = 2.25, $p = .046$, $\eta_p^2 = .14$, 90% CI [.00, .20]: the AB amplitude (lag 1 vs. 6/7) was larger for complete (8%) compared to ungrouped configurations (4%), *t*(14) = 4.20, $p = .001$, $d = 1.09$, 95% CI [.43, 1.72]. Partially grouped configuration (5%) exhibited a marginal difference relative to complete configurations ($p = .067$), but did not differ from ungrouped configurations ($p = .61$). Thus, the AB was significantly modulated by grouping strength. However, importantly, this grouping modulation occurred in the reverse order compared to, for instance, Experiment 1, with the complete T2 configuration now leading to the strongest (rather than the smallest) AB.

T1 accuracy. The mean percentage of correct responses for T1 was 97%. A two-way repeated measures ANOVA with the factors T2-target type (complete, partially grouped, ungrouped) and lag (1, 2, 6, 7) revealed only a lag effect, *F*(3, 42) = 15.71, $p < .0001$, $\eta_p^2 = .53$, 90% CI [.31, .63]: As in the previous experiments, T1 performance was somewhat reduced at lag 1 (95%; *p*s < .005), while being comparable for lags 2, 6, and 7 (98%, 98%, and 99%, respectively; *p*s > .83). No main or interaction effect involving T2-target type was revealed (*F*s < 1.4, *p*s > .23).

Figure 7 about there

Discussion

Experiment 3 showed overall a somewhat higher level of performance (possibly due to the change of task), but nevertheless again demonstrated a graded effect of T2 grouping on the AB, indicating that, especially at short lags, discrimination of the arrow target (orientation) was substantially influenced by the surrounding, task-irrelevant object configuration. Thus, grouping does modulate performance, in particular when resources are occupied by T1-related processing. It has been shown that a physically salient T1 stimulus engenders a reduction in performance at short lags even when there is no need to attend to T1 (Raymond et al., 1992). In the present study, we observed a modulation by the task-irrelevant grouping at short lags, which suggests the transient reduction in performance is not only owing to salient features (e.g. the red color) of the T1 object but also dependent on the irrelevant T2 grouping structure.

However, in contrast to Experiments 1 and 2, the effect of T2 configuration was *reversed*, with the smallest AB for ungrouped, an intermediate AB for partially grouped, and the largest AB for complete T2 configurations. This reversed AB pattern suggests that grouping, rather than being beneficial for the arrow discrimination task, did actually impair performance. An explanation for this pattern might be that attention was automatically captured by the task-irrelevant complete-object configuration, and as a result discrimination of the target orientation was hampered. Ungrouped T2 configurations, by contrast, attracted attention less and, consequently, more resources were available for the effective discrimination of the arrow target. In addition, unlike in the previous two experiments, T2 performance clearly reached the level of T1 accuracy at (or actually, well before) lag 7 (i.e., after 840 ms), for all types of configuration. This indicates that a task-irrelevant grouping may influence the efficiency of attentional target selection, thus modulating the AB primarily at short lags. However, the fact that this modulation was rather transient suggests that, in Experiment 3, grouping did not affect short-term memory consolidation (i.e., post-selective, stage-2 processing) of the T2 target.

Of note, Experiment 3 did also not reveal evidence of lag-1 sparing. This might again be due to a rather long lag (i.e., > 100 ms, see Olivers & Meeter, 2008, and Potter et al., 2002) and because of category and location switches that occurred between the two targets presented, which have previously been shown to hamper T2 processing (Kawahara et al., 2003; Di Lollo et al., 2005; Visser et al., 1999).

General Discussion

The AB phenomenon demonstrates that the human visual system is limited in its ability to extract durable mental representations from the rapidly changing, continuous flow of information across time. The present study investigated whether the AB effect is modulated by perceptual grouping in Kanizsa subjective figures, using a dual-target RSVP paradigm – the aim being to determine how attention is allocated to more or less structured visual information over time. Consistent with our predictions, the results showed that the AB effect is strongly modulated by T2 grouping strength: In Experiment 1, complete T2 groupings resulted in a smaller AB and in an increased overall performance compared to ungrouped (control) stimuli that consisted of the same pacman inducers which did, however, not induce an integrated percept. Experiment 2 replicated this pattern of results and further showed that the benefit of grouping in T2 can arise irrespective of the perceptual structure in T1 (Experiments 1 vs. 2). Finally, in Experiment 3, a modulation of grouping in T2 was obtained even though the task was entirely unrelated to the object configurations. In contrast to Experiment 1, performance in Experiment 3 revealed the largest AB when a T2 target was presented concurrently with a complete-object configuration. Together, this pattern of results suggests that identical inducer elements may differ in the extent to which an emergent global object is formed, which in turn affects the magnitude of the AB. Thus, grouping of separate parts into a coherent whole either attenuates or enhances the AB, depending on whether grouping is relevant or irrelevant to performing the task (Experiments 1 vs. 3). Overall, our results indicate that temporal attention is modulated by emergent objects.

Grouping modulates temporal object processing

Why does grouping in T2 modulate the allocation of attention in time? According to the two-stage model (Chun & Potter, 1995), after initial perceptual processing of the incoming sensory information, the perceptual representation must be encoded in a capacity-limited short-term memory system to ensure a stable and durable representation until a response can be issued. If this consolidation process is not accomplished, the perceptually processed item is ephemeral and rapidly overwritten by the items that appear subsequently in the RSVP stream. In this view, the AB reflects a post-perceptual, attentional mechanism of limited processing capacity, which subserves the consolidation of items into working memory (Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998; Vogel et al., 1998). With salient items – for instance, grouped objects such as Kanizsa figures (Rauschenberger & Yantis, 2001; Senkowski et al., 2005) – a processing advantage should arise relatively early, at the initial stage of perceptual coding, with the global structure of grouped objects allowing for more efficient detection compared to ungrouped configurations that lack a global representation (e.g., Conci et al., 2007a, 2009). That is, pre-attentive grouping would generate a salient structure that is more resistant to temporal decay at stage 1 – which would permit the global (structured) object to more efficiently consolidated at the subsequent, capacity-limited processing stage, effectively reducing the amount of interference in the AB. Thus, as a result of rather efficient and fast processing of a grouped T2, consolidation at stage 2 can commence earlier and proceed faster, as compared to a less structured T2, in turn facilitating the maintenance of the grouped object in working memory (see, e.g., Chen et al., 2016; Chen, Töllner, Müller, & Conci, 2018, for a related finding). In support of this view, Experiments 1 and 2 consistently showed overall superior performance for grouped than for ungrouped T2s, even at longer lags when T2 processing was no longer affected by T1 processing. This sustained difference indicates that retaining an item in memory is influenced by the object structure. Moreover, the performance difference for grouped (vs. ungrouped) T2s was largest at short intervals (in all experiments), where capacity-limited resources were most likely occupied by processes relating to T1. This further shows that, in addition, attentional limitations imposed by the AB can be overcome, to a significant extent, by grouping in the target, making processing more robust and more efficient in face of the lack of limited-capacity resources (Experiments 1 and 2).

However, it should be noted that – although grouping likely increased the coding efficiency of complete-object targets (i.e., it enabled consolidation to begin earlier and to require fewer attentional resources), which manifested in an attenuated AB (Experiments 1 and 2) – our results nevertheless revealed a clear AB for all grouping types. This might be taken to suggest that grouping of disparate items into a coherent whole nevertheless requires a certain amount of attentional resources in order to select and retain a relevant target item until the response is issued (Joseph et al., 1997; Braun, 1998; see also Gögler, Finke, Keller, Müller, & Conci, 2016; Conci, Groß, Keller, Müller, & Finke, 2018). However, increasing the efficiency with which the stimulus is encoded (e.g., by inducing grouping) in turn reduces the attentional load and, consequently, reduces the AB (see also Braun, 1998; Joseph, Chun, & Nakayama, 1998).

Additional support for an early-processing account of grouped objects derive from the results of Experiment 2, in which T1 presented a partially grouped (triangle) object that was more effective in binding attentional resources than the ungrouped T1 in Experiment 1. While the global T1 triangle in Experiment 2 led – at least to some extent - to an overall increased AB effect (as compared to the local T1 configuration in Experiment 1), the modulation of grouping in T2 was unaffected by this change in T1. This further supports the view that the benefit of grouping occurs because salient perceptual structures by themselves allow for a more efficient encoding of the grouped configurations (rather than arising from some top-down mediated sharing of resources between T1 and T2). That is, grouping renders particularly stable perceptual representations that are resilient in the face of interfering stimulation when only limited resources are available.

Consistent with this view, in visual search tasks, Kanizsa figures can act as a (non-informative) spatial cue, or in terms of an attractor for spatial attention, that facilitates detection of a target appearing at the same, circumscribed location (Senkowski et al., 2005; Conci, Müller, & von Mühlenen, 2013). However, the results of the present Experiment 3 show that when a comparable setup is used in an AB paradigm, a cost associated with the grouped object is observed, rather than efficient cueing of attention to the arrow target. This may come about as a result of the rapid succession of the stimuli in the RSVP stream. The Kanisza-type configuration may act as a salient distractor (i.e., it may capture attention), from which attention must be disengaged for the system to become able to discriminate the task-relevant arrow stimulus. However, by the time this is accomplished, the (briefly presented) target has already disappeared – resulting in a performance cost and in an increased AB. Of note, the task-irrelevant grouping modulated the detection of T2 primarily at short lags, whereas at longer lags T2 performance reached the same level as T1 performance, for all grouping types (complete, partially grouped, and ungrouped). This pattern contrasts with Experiments 1 and 2, in which (in these experiments) the task-relevant Kanizsa grouping not only modulated the immediate allocation of attention, but also the subsequent short-term memory consolidation of T2 at longer lags. This illustrates that task-irrelevant groupings can generate transient costs, whereas task-relevant groupings can yield sustained benefits – where the latter effect is likely owing to the encoding-into-memory of the (more or less grouped) task-relevant items.

Representing higher-order object files

When processing multiple objects in rapid succession, a key requirement of the visual system is its ability to select and consolidate potentially relevant information into an enduring representation, referred to as an "object file" (Kahnemann & Treisman, 1984). Raymond (2003) proposed that the creation of a new object file plays a key role in triggering the AB (see also Kellie & Shapiro, 2004; Conci & Müller, 2009). In line with such an object file account, we observed superior performance for T2 identification and an attenuated AB when T2 was identical in shape to T1 (see Experiment 2). Since an object file has already been set up upon the presentation of T1, with a same-object T2, the identical object file needs only to be updated – as a result of which the AB is reduced. In addition, integration might arise when two targets provide complementary shapes in close temporal proximity, as evidenced by a significant drop in performance across lags for partially grouped T2s (i.e., when T1 and T2 present Kanizsa triangles of opposite orientations; see Experiment 2). In this case, a “star” representation was more likely reported for T2, indicative of some form of misbinding across T1 and T2. These findings support an integration account as proposed by Hommel and Akyürek (2005), which assumes that it is difficult to segregate a continuous, rapid stream of visual information into discrete events. In this view, the closer in time two pieces of information appear, the more likely they are integrated into the same episodic trace – a finding which has been demonstrated using various types of objects and groupings (Bowman & Wyble, 2007; Karabay & Akyürek, 2017).

Conclusion

Whereas perceptual grouping can modulate the allocation of selective attention across visual space, the present findings show that structures provided by grouping can also influence the processing of targets in time. For instance, grouped targets lead to overall enhanced performance and a reduced AB effect, where the benefits from grouping are sustained, suggesting that they arise at an early, perceptual locus prior to attentional selection, thus facilitating both the detection of integrated structures and their subsequent consolidation into an enduring object file in working memory. By contrast, grouping in task-irrelevant items can transiently impair concurrent target processing, where this cost (from complete-object distractors) may be attributed to attentional capture, hindering efficient selection of the target. Together, this set of findings shows that grouping can substantially modulate the processing of objects in time.

Acknowledgements

This work was supported by project grants from the German Research Foundation (DFG; FOR 2293/1). Siyi Chen received a scholarship from the China Scholarship Council (CSC).

Open Practices

All data have been made publicly available via the open science framework and can be accessed at https://osf.io/we7aj/.

.

References

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General, 134*(2), 258-281.

Awh, E., Serences, J., Laurey, P., Dhaliwal, H., van der Jagt, T., & Dassonville, P. (2004). Evidence against a central bottleneck during the attentional blink: Multiple channels for configural and featural processing. *Cognitive Psychology, 48*(1), 95-126.

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*(1), 38-70.

Braun, J. (1998). Vision and attention: the role of training. *Nature, 393*(6684), 424-425.

Broadbent, D. E., & Broadbent, M. H. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics*, *42*(2), 105-113.

Brooks, J. L. (2015). Traditional and new principles of perceptual grouping. In J. Wagemans (Ed.), *The Oxford Handbook of Perceptual Organization* (pp. 57-87). Oxford University Press.

Chen, S., Glasauer, S., Müller, H. J., & Conci, M. (2018). Surface filling-in and contour interpolation contribute independently to Kanizsa figure formation. *Journal of Experimental Psychology: Human Perception and Performance, doi: 10.1037/xhp0000540.*

Chen, S., Müller, H. J., & Conci, M. (2016). Amodal completion in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(9), 1344-1353.

Chen, S., Töllner, T., Müller, H. J., & Conci, M. (2018). Object maintenance beyond their visible parts in working memory. *Journal of Neurophysiology*, *119*(1), 347–355.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109-127.

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.

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.

Conci, M., Groß, J., Keller, I., Müller, H. J., & Finke, K. (2018). Attention as the 'glue' for object integration in parietal extinction. *Cortex*, *101*, 60-72.

Conci, M., & Müller, H. J. (2009). The "beam of darkness": Spreading of the attentional blink within and between objects. *Attention, Perception, & Psychophysics*, *71*(8), 1725-1738.

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.

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.

Conci, M., Müller, H. J., & Elliott, M. A. (2009). The computation of shape orientation in search for Kanizsa figures. *Perception, 38*, 173-185.

Conci, M., Müller, H. J. & von Mühlenen, A. (2013). Object-based implicit learning in visual search: Perceptual segmentation constrains contextual cueing. *Journal of Vision, 13*(3):15, 1-17.

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.

Davis, G., & Driver, J. (1994). Parallel detection of Kanizsa subjective figures in the human visual system. *Nature*, *371*(6500), 791-793.

Di Lollo, V., Kawahara, J. I., Ghorashi, S. S., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control?. *Psychological research, 69*(3), 191-200.

Garner, W. R., & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior*, *2*(5), 446-452.

Gögler, N., Finke, K., Keller, I., Müller, H. J. & Conci, M. (2016). Object integration requires attention: Visual search for Kanizsa figures in parietal extinction. *Neuropsychologia, 92*, 42-50.

Hommel, B., & Akyürek, E. G. (2005). Lag-1 sparing in the attentional blink: Benefits and costs of integrating two events into a single episode. *The Quarterly Journal of Experimental Psychology Section A, 58*(8), 1415-1433.

Huang, Y. M , Baddeley, A., Young, A. W. (2008). Attentional capture by emotional stimuli is modulated by semantic processing. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 328-339.

Jackson, M. C., & Raymond, J. E. (2006). The role of attention and familiarity in face identification. *Perception & Psychophysics, 68*(4), 543-557.

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, *36*(2), 138-202.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a 'preattentive' feature search task. *Nature, 387*(6635), 805-807.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1998). Reply: Vision and attention: the role of training. *Nature, 393*(6684), 425-425.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. A. Davies (Eds.), *Varieties of Attention* (pp. 29-61). New York: Academic Press.

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.

Karabay, A., & Akyürek, E. G. (2017). The effects of Kanizsa contours on temporal integration and attention in rapid serial visual presentation. *Attention, Perception, & Psychophysics, 79*, 1742-1754.

Kawahara, J., Zuvic, S. M., Enns, J. T., & Di Lollo, V. (2003). Task switching mediates the attentional blink even without backward masking. *Perception & Psychophysics, 65*, 339-351.

Keil, A., & Ihssen, N. (2004). Identification facilitation for emotionally arousing verbs during the Attentional Blink. *Emotion, 4*(1), 23-35.

Keil, A., Ihssen, N., & Heim, S. (2006). Early cortical facilitation for emotionally arousing targets during the attentional blink. *BMC Biology, 4*(1), 23.

Kellie, F. J., & Shapiro, K. L. (2004). Object file continuity predicts attentional blink magnitude. *Perception & Psychophysics*, *66*(4), 692-712.

Koffka, K. (1935). *Principles of Gestalt Psychology.* New York: Harcourt.

Luck, S. J. (1998). Neurophysiology of selective attention. In H. Pashler (Ed.), *Attention* (pp. 257-295). London: UCL Press.

Maki, W. S., Bussard, G., Lopez, K., & Digby, B. (2003). Sources of interference in the attentional blink: Target-distractor similarity revisited. *Perception & Psychophysics*, *65*(2), 188-201.

Marini, F., & Marzi, C. A. (2016). Gestalt perceptual organization of visual stimuli captures attention automatically: Electrophysiological evidence. *Frontiers in Human Neuroscience*, 10, 446.

Martens, S., & Wyble, B. (2010). The attentional blink: Past, present, and future of a blind spot in perceptual awareness. *Neuroscience & Biobehavioral Reviews*, *34*(6), 947-957.

Moore, C. M., & Egeth, H. (1997). Perception without attention: evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(2), 339-352.

Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1171.

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.

Olivers, C. N., & Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychological Review, 115*(4), 836-863.

Potter, M. C., Staub, A., & O'Connor, D. H. (2002). The time course of competition for attention: attention is initially labile. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1149-1162.

Rauschenberger, R., & Yantis, S. (2001). Attentional capture by globally defined objects. *Perception & Psychophysics, 63*(7), 1250-1261.

Raymond, J. E. (2003). New objects, not new features, trigger the attentional blink. *Psychological Science*, *14*(1), 54-59.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink?. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849-860.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1995). Similarity determines the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 653-662.

Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: evidence for low-level grouping. *Psychological Review*, *102*(1), 101-130.

Senkowski, D., Röttger, S., Grimm, S., Foxe, J. J., & Herrmann, C. S. (2005). Kanizsa subjective figures capture visual spatial attention: evidence from electrophysiological and behavioral data. *Neuropsychologia, 43*(6), 872-886.

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1997). The attentional blink. *Trends in Cognitive Sciences*, *1*(8), 291-296.

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 357-371.

Stein, T., Zwickel, J., Ritter, J., Kitzmantel, M., Schneider, W. X. (2009). The effect of fearful faces on the attentional blink is task dependent. *Psychonomic Bulletin & Review 16,* 104–109.

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.

Visser, T. A., Bischof, W. F., & Di Lollo, V. (1999). Attentional switching in spatial and nonspatial domains: Evidence from the attentional blink. *Psychological Bulletin, 125*(4), 458-469.

Visser, T. A., Bischof, W. F., & Di Lollo, V. (2004). Rapid serial visual distraction: Task-irrelevant items can produce an attentional blink. *Perception & Psychophysics*, *66*(8), 1418-1432.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1656-1674.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychological Research, 4*(1), 301-350.

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.

Woodman, G. F., & Luck, S. J. (2003). Dissociations among attention, perception, and awareness during object-substitution masking. *Psychological Science, 14*(6), 605-611.

Figures

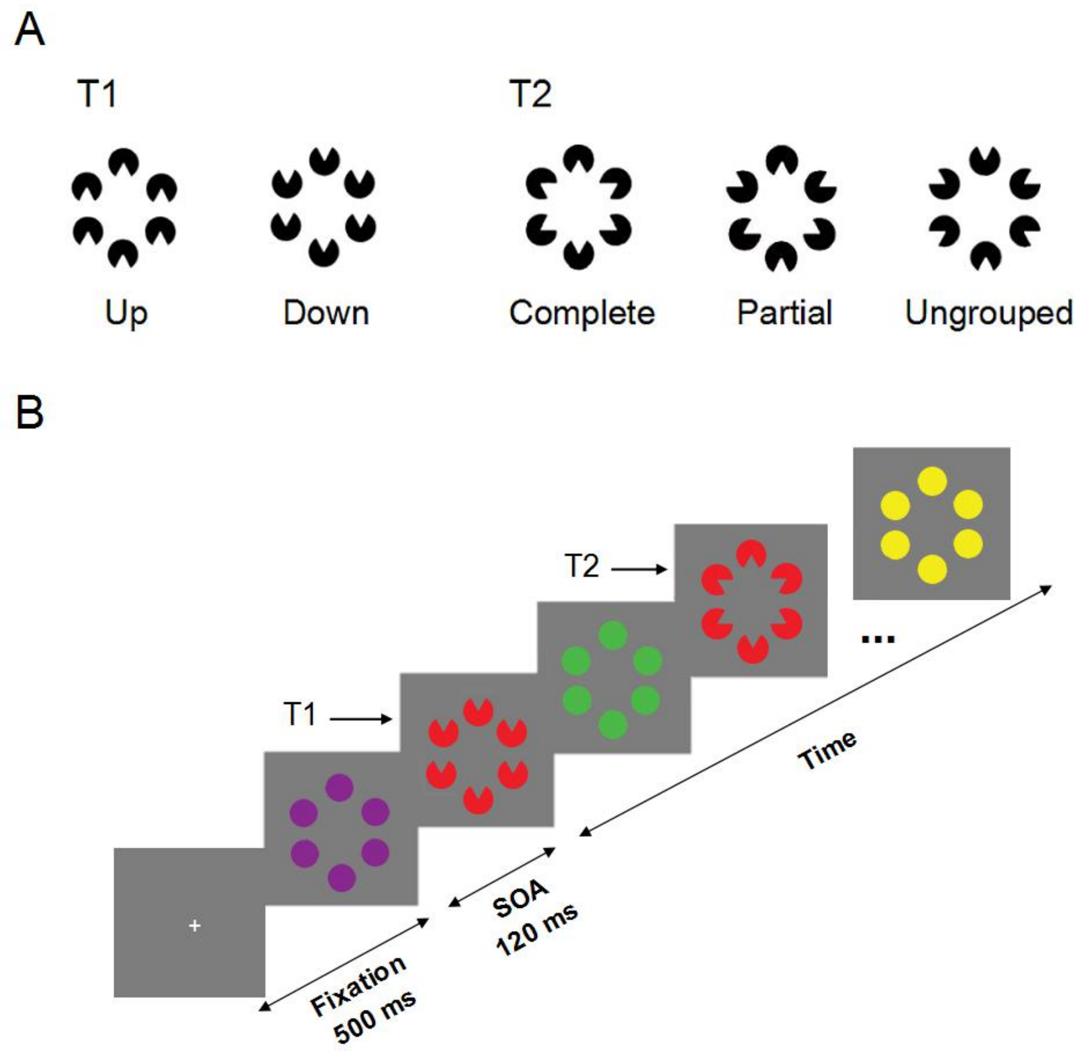

Figure 1. Stimuli and display sequence in Experiment 1. (A) Illustration of experimental stimuli for T1 (all pacman inducers oriented either up- or downwards) and T2 (left: complete; middle: partially grouped; right: ungrouped). (B) Schematic example of the RSVP sequence. Each trial presented a sequence of 10 displays, which consisted of either six complete distractor disks (non-red items) or the T1 and T2 target arrangements (red items).

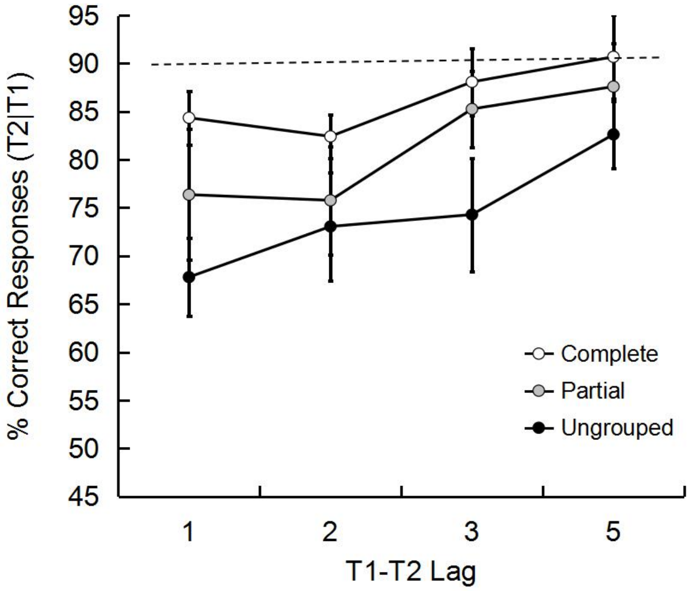

Figure 2. Mean percentage of correct identifications of T2 (given a correct T1 response) in Experiment 1. Correct identifications are presented as a function of the temporal lag from the onset of T1 to the onset of T2, separately for the different T2-target types (complete, partially grouped, and ungrouped configurations). The dashed horizontal line indicates the level of overall T1 accuracy. Error bars denote 95% within-subject confidence intervals.

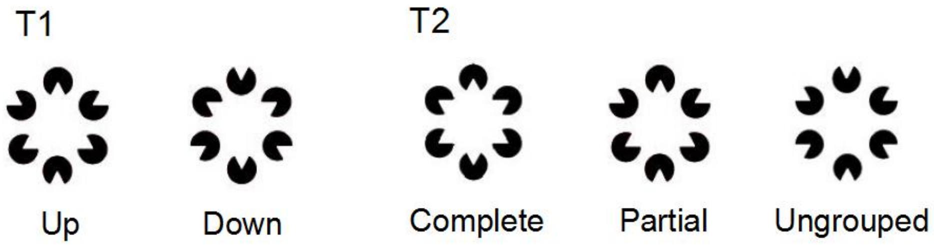

Figure 3. Example target configurations for T1 (up vs. downward pointing triangles) and T2 (complete, partially grouped, or ungrouped) in Experiment 2.

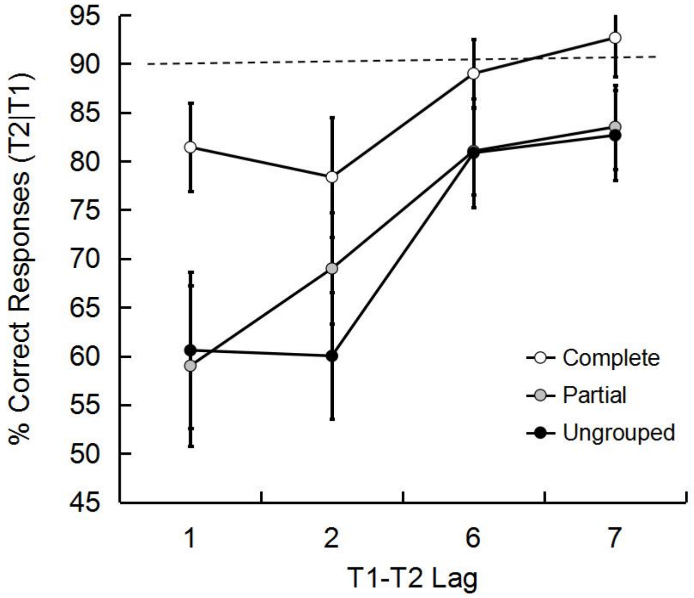

Figure 4. Mean percentage of correct identifications of T2 (given a correct T1 response) in Experiment 2. Correct identifications are presented as a function of the temporal lag from the onset of T1 to the onset of T2, separately for the different conditions (complete, partially grouped, and ungrouped configurations). The dashed horizontal line indicates the level of overall T1 accuracy. Error bars indicate 95% within-subject confidence intervals.

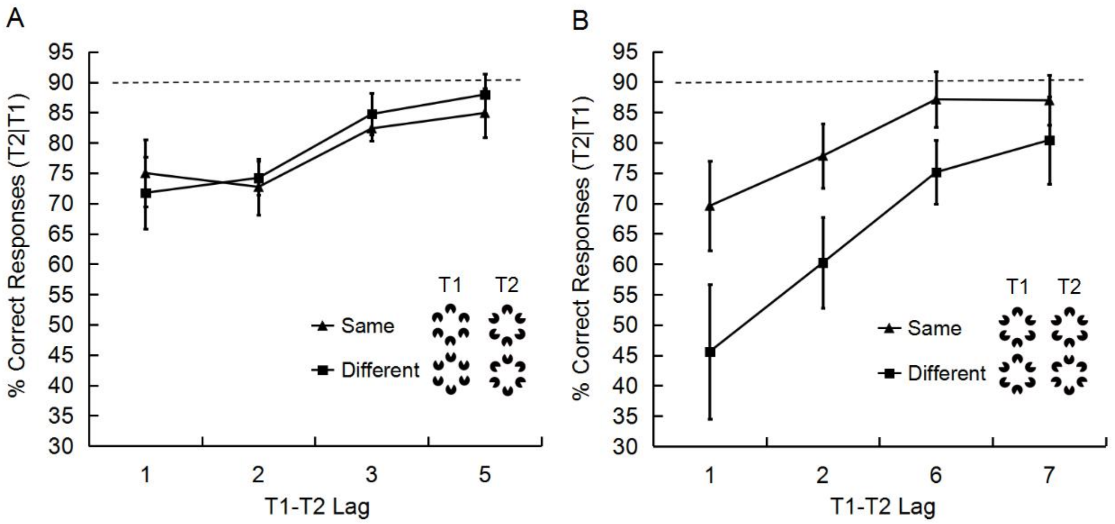

Figure 5. Mean percentage of correct identifications of T2 (given a correct T1 response) in Experiment 1 (A) and in Experiment 2 (B). Correct identifications are presented as a function of the temporal lag from the onset of T1 to the onset of T2, separately for same (matching) and different (mismatching) T1–T2 pointing directions (where T2 presented a Kanizsa triangle with up- or downward pointing direction). The dashed horizontal line indicates the level of overall T1 accuracy. Error bars indicate 95% within-subject confidence intervals.

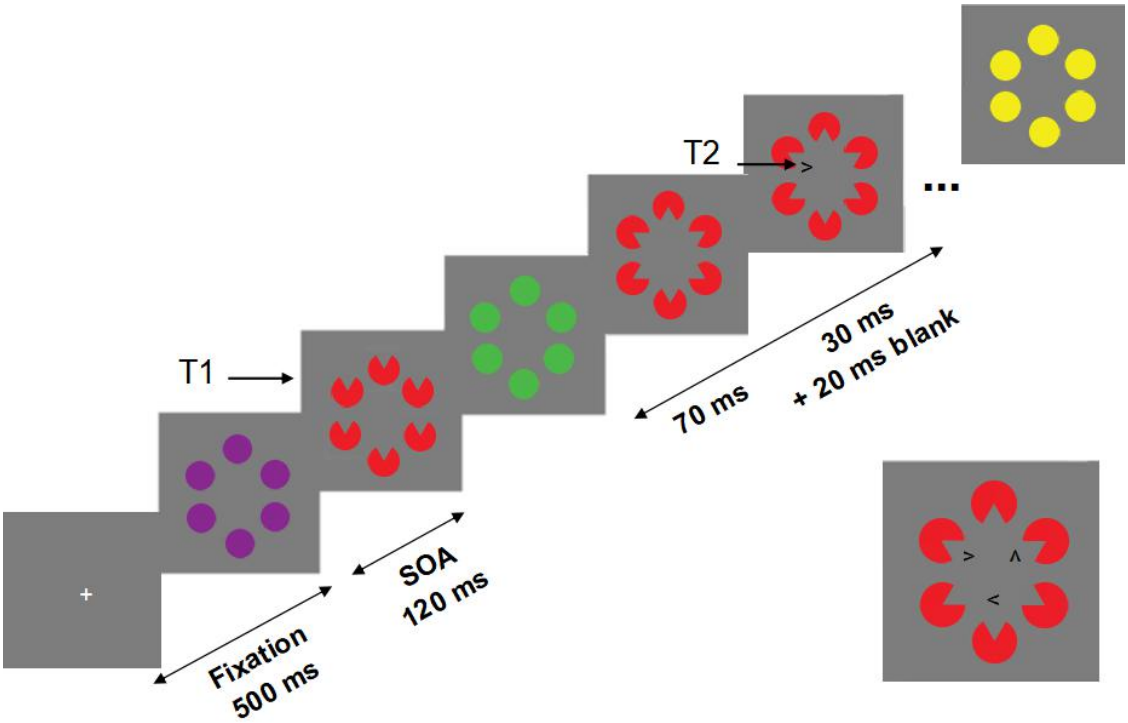

Figure 6. Schematic example of the RSVP sequence in Experiment 3. Each trial presented a sequence of 12 displays, which consisted of either six complete distractor disks (non-red items) or the T1 and T2 target. For T2, a complete, partially grouped, or an ungrouped configuration was presented (as in Experiment 1), but with an additional target arrow (i.e., an oriented “>”-sign) added to the display. Note that the T2 task was related only to the arrow (but not in any way to the grouping as presented in the Kanizsa-type configurations). The bottom right panel illustrates the three possible locations of the target arrow.

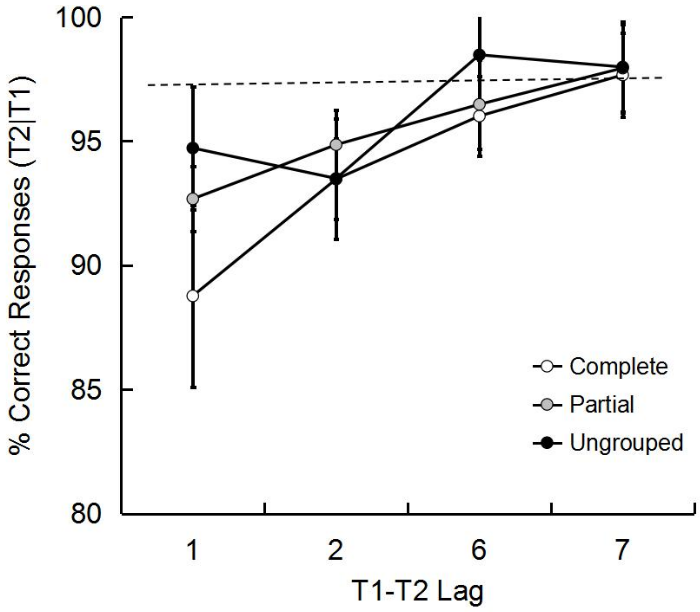

Figure 7. Mean percentage of correct identifications of T2 (given a correct T1 response) in Experiment 3. Correct identifications are presented as a function of the temporal lag from the onset of T1 to the onset of T2, separately for the different T2-target conditions (complete, partially grouped, and ungrouped configurations). The dashed horizontal line indicates the level of overall T1 accuracy. Error bars denote 95% within-subject confidence intervals.