Two's company, three's a crowd: Object individuation and recognition rely on common mechanisms

Ramakrishna Chakravarthi[a, 1] and Amy Herbert[a]

a. School of Psychology, University of Aberdeen, Kings College, Aberdeen, AB24 3FX UK

1. Corresponding author: rama@abdn.ac.uk

Highlights

- The relationship between object individuation and identification was tested
- Crowding (an identification process) impaired subitizing (an individuation process)
- Individuation is necessary for recognition, and is impaired in crowding
- Crowding and individuation share a common bottleneck: likely selective attention

Abstract

Object recognition is essential for navigating the real world. Despite decades of research on this topic, the processing steps necessary for recognition remain unclear. In this study, we examined the necessity and role of individuation, the ability to select a small number of spatially distinct objects irrespective of their identity, in the recognition process. More specifically, we tested if the ability to rapidly individuate and enumerate a small number of objects (subitizing) can be impaired by crowding. Crowding is flanker-induced interference that specifically impedes the recognition process. We found that subitizing is impaired when objects are close to each other (Experiment 1), and if the target objects are surrounded by irrelevant but perceptually similar flankers (Experiments 2-4). This impairment cannot be attributed to confusion between targets and flankers, wherein flankers are inadvertently included in or targets are excluded from enumeration (Experiments 3-4). Importantly, the flanker induced interference was comparable in both subitizing and crowding tasks (Experiment 4), suggesting that individuation and identification share a common processing pathway. We conclude that individuation is an essential stage in the object recognition pipeline and argue for a cohesive proposal that both crowding and subitizing are due to limitations of selective attention.

Keywords: Individuation; Recognition; Subitizing; Crowding; Enumeration; Attention

1. Introduction[1]

Recognising objects is a central function of the visual system. Over the last several decades, there has been extensive research on the mechanisms underlying object recognition (Biederman, 1987; DiCarlo, Zoccolan, & Rust, 2012; Marr, 1982; Ullman, 1996, 2007). Distilling and simplifying a substantial amount of this research, we might surmise that certain processing stages, such as feature detection, segmentation or individuation, and feature integration, are crucial for recognising objects. Nevertheless, there is no consensus regarding the necessity of these stages and their sequence in the object recognition pipeline. In a step towards a better understanding of the process, in this study we will examine the role of individuation in the object processing stream.

Spatial individuation, or selecting an object based on its location, irrespective of its identity, (see Mazza & Caramazza, 2015 for a review) has been argued to be an important step in object recognition (Trick & Pylyshyn, 1994; Xu & Chun, 2009). In addition, this ability forms one of the bases of numerical cognition (Gallistel & Gelman, 1992, 2000). It appears to be a necessary step for non-symbolically representing numbers and in apprehending numerosity (Piazza & Eger, 2016). Individuation and numerosity are thought to be primarily processed in the parietal cortex (Nieder, 2005; Xu & Chun, 2009). On the other hand, recognition is often considered to be computed in the lateral occipital and inferior temporal cortices (e.g., DiCarlo et al., 2012; Grill-Spector & Sayres, 2008; Tsao & Livingstone, 2008). This apparent discrepancy brings into sharp focus the debate about the role of individuation in recognition. Indeed, enumeration, individuation and recognition have rarely been examined together. Separate studies on these disparate capacities have led to roughly comparable yet differing conclusions about the steps required for recognising objects and their precise sequence.

1.1. Processing pipeline for object recognition

1.1.1. Individuation studies

Individuation is often assessed using tasks such as multiple object tracking and subitizing (Trick & Pylyshyn, 1994). These tasks often eschew the requirement to identify object(s) but require participants to track the positions of identical objects or to enumerate them. This

---

[1] Data collected in this study are available at: http://dx.doi.org/10.20392/165cee3b-5d4b-4945-829e-cf07ee222ac0

approach is expected to isolate processes specific to individuation. Such studies have demonstrated that humans can individuate about 3-4 objects at a time. Findings from these studies (e.g., Trick & Pylyshyn, 1994; Xu & Chun, 2009) have been taken to advocate for a sequence of stages in the object recognition process. First, object features are independently registered by the visual system and in parallel. These features are then segmented by grouping and figure-ground segregation processes. This is the individuation stage, where the objects are indexed or tagged by the visual system and their features are clustered together, but their identities are still unknown. There is an ongoing debate regarding whether attention is required for the operation of this stage (Mazza & Caramazza, 2015; Trick & Pylyshyn, 1994; Vetter, Butterworth, & Bahrami, 2008). Nevertheless, it is generally agreed that up to four objects can be individuated at any one time. These individuated objects are then selected for further processing by attention where their features are integrated. This integrated representation is subsequently recognised. According to this line of reasoning, individuation is an integral part of the recognition process and has limited resources. These steps are illustrated in Fig 1.

A specific example of an object recognition pipeline that includes individuation is Xu and Chun’s (2009) ‘Neural Object-File Theory’, according to which, at the individuation stage, up to four objects can be selected at once by attention, regardless of their complexity. These objects are coarsely represented with the features in an unbound state. The limitation of this stage restricts the range of efficiently enumerable objects to four, an ability known as subitizing (Jevons, 1897; Kaufman, Lord, Reese, & Volkmann, 1949).The features of these individuated objects are then bound together into a coherent representation at the integration stage; such objects are represented with high fidelity. These representations are processed by downstream regions (e.g., temporal cortex) to determine their identity.

1.1.2. Visual crowding studies

A successful approach to studying object recognition has been to examine conditions where it fails. One such situation is visual crowding, where recognition of an object is impaired when it is surrounded by similar clutter (Bouma, 1970; Levi, 2008; Pelli & Tillman, 2007). It has been shown that crowding does not impair object detection, but only affects identification (Levi, Hariharan, & Klein, 2002; Pelli, Palomares, & Majaj, 2004). Further, a crowded and hence unidentifiable grating can yet lead to orientation (He, Cavanagh, &

Intriligator, 1996) and motion (Whitney, 2005) aftereffects. That is, crowding selectively affects feature binding and identification without altering prior processes. These findings have been taken to suggest that crowding is a mid-level processing failure (Chakravarthi & Cavanagh, 2009; Shin, Chung, & Tjan, 2017).

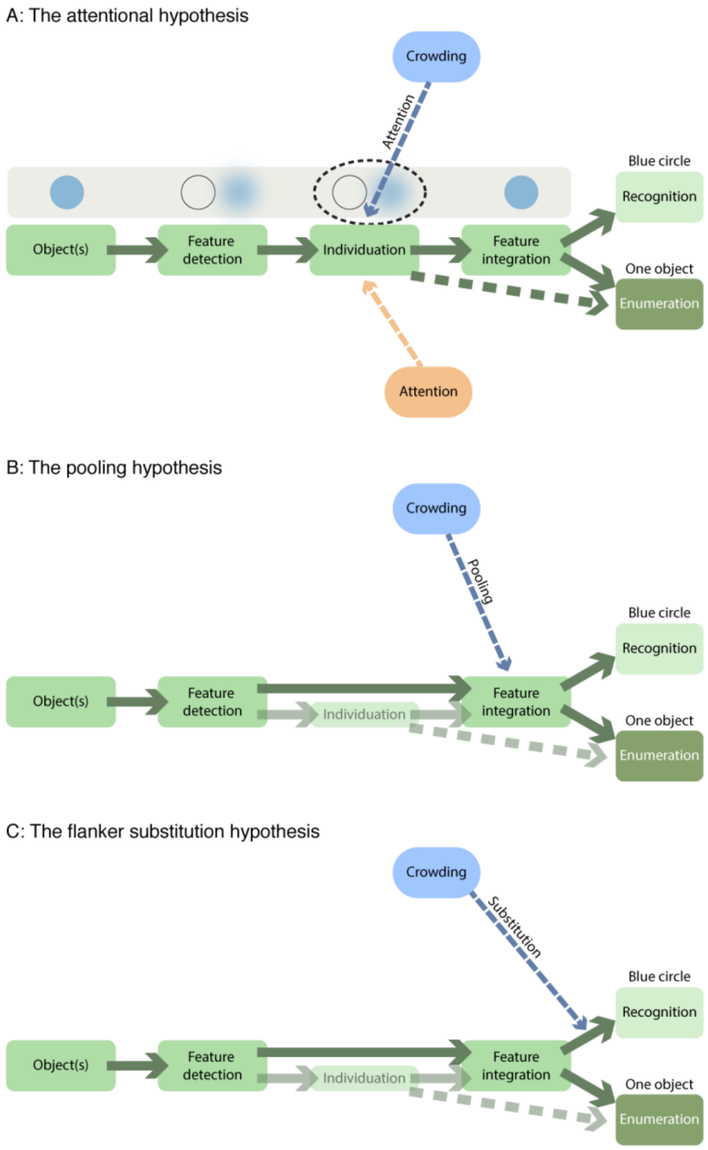

Figure 1: Processing pipeline for object recognition according to various theories of crowding. **A.** *The attentional hypothesis*. The flowchart (in green) depicts the processing pipeline for object recognition. The ‘events’ in the grey strip above the flowchart illustrate the stages in the pipeline. First, the features of an object are detected independently and in parallel. Next, these registered features are individuated and indexed by attention. At this stage, the features are segmented and clustered but remain unbound. The output of this stage might be sufficient for the enumeration task (thick dashed green arrow). Features are then bound together at the feature integration stage. The bound representation is then used by downstream processes such as recognition and enumeration. When a

single object is present, this process occurs smoothly without interference. If multiple objects are close to each other, then their individuation is impaired leading to crowding (He et al., 1996; Intriligator & Cavanagh, 2001). According to this hypothesis, surrounding flankers should affect individuation and hence subitizing. **B.** *The pooling hypothesis*, on the other hand, does not explicitly require an individuation stage, although it can potentially be included, in principle (faded parts of the pathway). Here, the detected features are integrated at the second stage of integration. Crowding occurs due to inappropriate pooling at this stage. The output of feature integration is then used for recognition and enumeration. However, this hypothesis does not exclude the possibility that the enumeration (and hence the individuation step) pathway is distinct from the recognition pathway. In either case, this hypothesis predicts that subitizing is not impaired by crowding (see section 1.1.3.2. for details). **C.** *The flanker substitution hypothesis* is similar to the pooling hypothesis, except that crowding occurs after feature integration, by the swapping of intact targets and flankers. It does not affect subitizing.

There are many accounts of crowding. A commonly held ‘pooling’ view is that it is the consequence of inappropriate integration of features that belong to distinct but closely spaced objects (Levi, 2008; Pelli et al., 2004). This inappropriate integration can take the form of averaging of features (Greenwood, Bex, & Dakin, 2009; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). It has also been posited that features can migrate or be swapped between objects (Nandy & Tjan, 2007). A second ’flanker substitution’ account posits that crowding occurs due to the loss of position information and hence observers mistakenly report one of the flanking objects (Ester, Klee, & Awh, 2014; Ester, Zilber, & Serences, 2015; Strasburger & Malania, 2013). In either case, the findings have been argued to support a simple two stage model of object recognition (Fig 1B-C). The first step is independent and parallel detection of object features across the visual field. The second step involves the integration of these features into representations that are recognised by downstream processes. If two or more objects are close to each other, then a) their features are ‘pooled’, averaged, or swapped, or b) their features are appropriately integrated, but during the post-integration stage their position information is lost and whole objects are swapped, leading to crowding. Hence crowding is a failure at the stage of feature integration or later. Note that this simple model does not explicitly include object individuation as a processing stage.

A third, ‘attentional’ account of crowding argues that crowding is due to the limitation of attentional resolution (He et al., 1996; Intriligator & Cavanagh, 2001). That is, when multiple objects are close to each other, selective attention cannot isolate and select a single object.

It therefore inadvertently selects multiple objects resulting in an inability to resolve and identify the target object. According to this proposal, once features are detected, there is a step of individuation, where clusters of features are selected by attention, which is then followed by feature integration. Crowding, here, is a failure at the stage of individuation. Intriligator and Cavanagh (2001) tested this proposal in a study on attentional resolution, where they presented participants with uniformly spaced discs in the periphery and asked them to start at one randomly selected disc and then mentally 'step' across them one at a time according to verbal instructions (e.g., left or right). Crucially, this task does not require participants to identify the objects but to individuate them. They found that the more densely packed the discs were, the more difficult the participants found to step across discs as instructed. The minimal inter-disc spacing at which impairment in performance was no longer observed matched the distance estimated for unimpaired identification in standard crowding tasks, where participants are asked to identify a flanked target. The authors therefore concluded that attentional selection and hence individuation is impaired when objects are too closely spaced. They argued that this underlies the failure of identification in crowding. Note that while the findings from the Intriligator and Cavanagh (2001) study show that the spatial constraints on individuation are comparable to those observed in crowding, they do not directly test whether crowding occurs at the stage of individuation. That conclusion is inferred from the similarity of findings across the two tasks. A direct test of crowding would include determining if irrelevant flankers impair individuation, just as they would identification. A further stringent test would be to determine if this impairment is the same for both individuation and identification tasks performed on the same stimulus.

One might therefore surmise that individuation is an essential stage in the object recognition process. Nevertheless, current computational models of crowding (e.g., Harrison & Bex, 2015; Keshvari & Rosenholtz, 2016; van den Berg, Roerdink, & Cornelissen, 2010) do not incorporate a stage of object individuation before their features are pooled. This might be one reason why such models have difficulty explaining the results of studies where object-level grouping between target and flankers or amongst flankers affects target identification performance (e.g., Herzog, Sayim, Chicherov, & Manassi, 2015). Models that explicitly or implicitly involve segmentation of feature sets have been shown to capture these grouping effects (Chaney, Fischer, & Whitney, 2014; Francis, Manassi, & Herzog, 2017).

1.1.3. Subitizing and crowding

In the current study, we will focus on subitizing as an index of individuation. Although alternative theories exist to account for subitizing (e.g., pattern recognition: Krajcsi, Szabó, & Mórocz, 2013; Logan & Zbrodoff, 2003; Mandler & Shebo, 1982; estimation process: Balakrishnan & Ashby, 1992; Dehaene & Changeux, 1993; Gallistel & Gelman, 1991), there is substantial evidence that subitizing is subserved by individuation (Franconeri, Bemis, & Alvarez, 2009; Mazza & Caramazza, 2015; Xu & Chun, 2009). That is, objects need to be individuated in order to be subitized. However, it should be noted that there is no consensus regarding whether the subsequent feature integration stage is necessary for enumeration and subitizing (e.g., Xu & Chun, 2009). It is possible that the output of the individuation stage is sufficient for successful subitizing (thick dashed green arrow in Fig 1A).

The two sets of studies, on crowding and individuation, provide a mixed picture regarding the role of individuation in object recognition (Fig 1). The current study is designed to shed light on their relationship by testing if crowding impacts the individuation stage. If it does, we argue that the stage of individuation must be incorporated into models of crowding and hence object recognition. The results will also specify the mechanism underlying crowding. Below, we will work through the predictions of the various theories of crowding for the outcome of a study testing if subitizing can be crowded, keeping in mind the processing pipeline illustrated in Fig 1.

1.1.3.1. Attentional hypothesis

If crowding occurs at the individuation stage, as proposed by the attentional hypothesis, then subitizing should be impaired by the presence of flankers. In addition to supporting the notion that crowding is a consequence of attentional limitations, this outcome would imply that a stage of individuation is necessary for object recognition (Fig 1A).

1.1.3.2. Pooling hypothesis

If crowding occurs due to feature pooling or averaging, then the prediction is not straightforward. Averaging of features (colour, orientation) by itself should not alter the number of perceived objects. Hence, we would not expect crowding to impair subitizing.

But, it could be argued that object positions are features too and these positions might be averaged or pooled (Greenwood et al., 2009). During the pooling process, the target's features are 'assimilated' towards those of the flankers (Greenwood et al., 2009; Mareschal, Morgan, & Solomon, 2010). That is, the target's perceived feature value is a weighted average of all features within a region of space around the target (Greenwood et al., 2009; Harrison & Bex, 2015; Parkes et al., 2001; van den Berg et al., 2010). Applying this logic to position information, we can expect that target locations are assimilated towards flanker positions and vice versa. Early reports of the phenomenology of crowding describe something like this being observed. Korte (1923) reported that observers described a crowded array of letters as "… if there is a pressure on both sides of the word that tends to compress it." (see Fig 2 in Tyler & Likova, 2007 for an illustration). In such a situation, numerosity can be underreported if the assimilated locations are closer than the visual system's ability to resolve objects. That is, if the perceived locations are too close, the visual system cannot separate the two objects and hence underestimates the numerosity. However, since visual acuity (two-dot resolution) is an order of magnitude finer than the inter-object distances used in typical crowding experiments (Anstis, 1974; Intriligator & Cavanagh, 2001), we think that it is not very likely that the visual system will be unable to resolve the pooled locations. Some recent observations support the idea that pooled locations might still be separable. Sayim and Wagemans (2017) have noted that, under crowded conditions, participants most often report omissions of individual features of objects. For example, they might not report one of the strokes in the letter A. However, observers don't seem to lose an entire object (or perceive an additional object). Summarising these findings, it appears that the averaging hypothesis predicts that crowding compresses objects together. This might potentially impair subitizing, although we think that this is unlikely, given the visual system's relatively high sensitivity in resolving objects.

1.1.3.3. Substitution hypothesis

If crowding is predominantly due to flanker substitution, where intact target and flankers are swapped, then enumeration and subitizing should remain unimpaired. Similarly, if crowding occurs due to feature migration or swapping of features, determining the number of objects should not be affected (Fig 1C).

1.1.3.4. Predictions and implications

To summarise, the attentional resolution hypothesis predicts that subitizing will be impaired by flankers and most of the other crowding hypotheses predict little to no effect of flankers on subitizing. Importantly, an impairment of subitizing by flankers would strongly suggest that an individuation stage should be included in crowding/object recognition models.

The impairment of subitizing by crowding would support the attentional hypothesis of crowding and individuation. It would commit us to including an individuation stage in the processing pipeline (Fig 1A). On the other hand, if crowding does not affect subitizing, it would strongly support the exclusion of attention as a mechanism of crowding (Fig 1B-C). That is, it would lend evidence against the proposal that crowding is an impairment of individuation. It would imply that one of the other proposed mechanisms (integration, averaging or substitution) is at play. Further, it would support the contention that individuation is not necessary for object recognition, although its involvement cannot be completely ruled out (faded pathways in Fig 1B-C).

2. Experiment 1: Internal crowding of subitizing

Crowding depends on the spacing between a target and its flankers. Crowding is eliminated if the target-flanker spacing exceeds approximately a) half the target's eccentricity when the target and flankers are aligned radially relative to fixation, or b) a quarter of target's eccentricity when flankers are aligned tangentially (Bouma, 1970; Toet & Levi, 1992). Following Pelli and Tillman (2008), we refer to this limit as Bouma's bound. Only objects within Bouma's bound crowd each other. However, it is important to keep in mind that the stated bounds are only rules of thumb and there is considerable variability across individuals (Toet & Levi, 1992). There are also exceptions to the rule (Herzog et al., 2015; Rosen, Chakravarthi, & Pelli, 2014), but these exceptions are not pertinent to the current study. Here, as a first step in testing the effect of crowding on subitizing, we investigated enumeration in the periphery as a function of inter-object distance.

2.1. Methods

2.1.1. Participants

Eighteen observers with normal or corrected to normal vision took part in this experiment. The first author took part in all experiments. All participants in this and subsequent experiments provided written informed consent. These experiments were approved by the Psychology Ethics Committee at the University of Aberdeen.

2.1.2. Materials

Stimuli were generated and displayed using MATLAB with Psychophysics Toolbox extensions (Kleiner, Brainard, & Pelli, 2007) on a 19” CRT screen with a resolution of 1024 x 768 pixels and a refresh rate of 100 Hz. The monitor was placed 50 cm from the observer, and the head was stabilised with a chin rest.

2.1.3. Stimuli and Procedure

1-6 target square ‘dots’ were presented on an isoeccentric circle of radius 10 deg centred on fixation (Fig. 2A). Four possible inter-dot distances were used, which we report in terms of geometric angular separation measured from fixation (θ): adjacent dots could be separated by 5^o, 10^o, 20^o or 40^o, equivalent to straight-line inter-dot spacing of 0.9, 1.7, 3.5, or 6.4 deg, respectively. To achieve this, we divided an imaginary circle of radius 10 deg around fixation into 72 locations, with adjacent locations separated by 5^o. Of these 72 locations, one location was randomly chosen on each trial, where the first object was placed. The remaining objects were placed at the appropriate distances from this location (gaps of 0, 1, 2, or 4 locations between adjacent dots). A jitter of ±2 pixels was applied both in the vertical and horizontal direction. At the tested eccentricity, an inter-dot spacing of less than 2.5 degrees should place the dots within crowding distance of each other. Note that no crowding is expected, for obvious reasons, when a single dot is presented; similarly, when two dots are presented, only weak crowding is expected (Pelli et al., 2004; Petrov & Meleshkevich, 2011). To ensure that participants were not simply assessing the total area occupied by the dots to determine numerosity, dot size varied from trial to trial and was randomly selected between 0.25 and 0.36 deg, but was the same within a trial. Target objects were black (2.1 cd/m^2) on a grey background (39.6 cd/m^2).

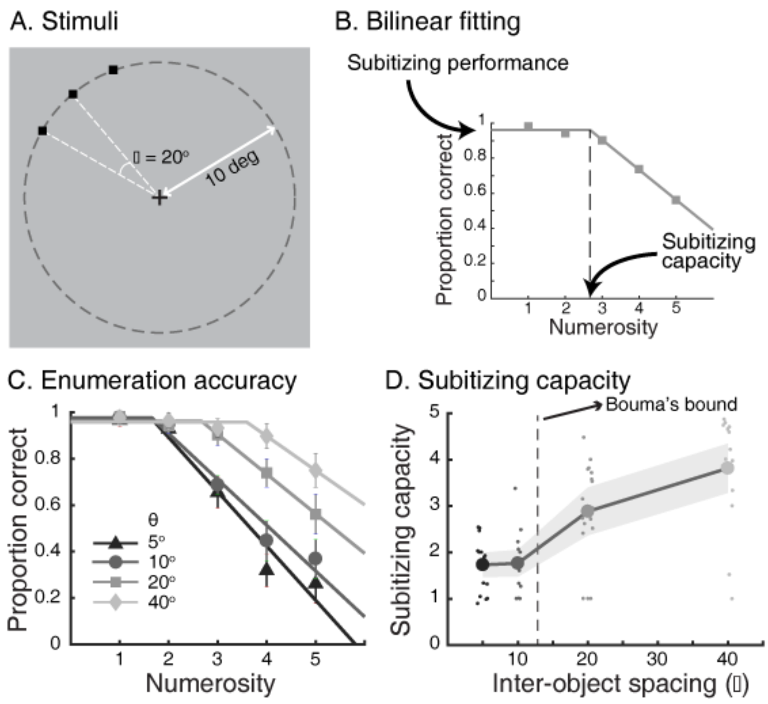

Figure 2: Stimulus setup and results of Experiment 1. **A**: 1-6 square dots were presented in the periphery on an imaginary circle (not shown in the experiment). The spacing between the dots (θ) was varied. **B**. Example of a bilinear fit. A bilinear function fits two straight lines to the data. The y-intercept of the flat line is the subitizing performance. The intersection of the two lines, indicated by a dashed line here, is the subitizing capacity. **C.** Accuracy of reporting the number of items for each spacing condition. Bilinear fits are also shown. Error bars are 95% CI. **D.** Subitizing capacity estimated for each inter-object spacing. Shaded areas are 95% CI. Each participant's capacity at each spacing is represented as a dot. The vertical dashed line is the classical Bouma's bound.

The order of spacing conditions was randomised within each block. Target dots were presented for 150 ms. Observers were asked to report, by means of the number pad on a keyboard, the number of dots while fixating a central cross within 1.5 s of stimulus onset. If they failed to respond, the trial was marked as incorrect. The next trial began 1 s after the response or 1 s after the response deadline passed. There were 40 trials per numerosity and spacing combination, with a total of 960 trials. No feedback was provided.

2.1.4. Data analysis: Estimating subitizing capacity and subitizing performance using bilinear fits

Following the usual practice in enumeration studies, data from trials with the highest numerosity were discarded because of the 'end effect'. The end effect is the better than expected performance for the highest presented numerosity, likely due to a bias for reporting the highest number in the presence of uncertainty (Piazza, Mechelli, Butterworth,

& Price, 2002). Bilinear functions were then fitted to the performance data. In experiments 1, 3 and 4 performance was measured in terms of accuracy and in Experiment 2, it was measured in terms of reaction times. A bilinear function fits two intersecting straight lines to the data: The first line had a fixed slope of 0, but the intercept was allowed to vary. This line indicates efficient enumeration of objects. Therefore, its intercept captures *subitizing performance*, which is the best performance across small numbers of items (see Fig 2B). The second line's slope and intercept were both allowed to vary. This line indicates inefficient or error-prone enumeration. The point of intersection, the breakpoint or the 'elbow', between these two lines indicates the *subitizing capacity*. This is the maximum number of objects that the participant can enumerate or individuate efficiently. We also assessed subitizing capacity while allowing the slope of the first line to vary (see Supplementary Results S1.2). The results were very similar as when the slope was fixed at zero, with only a small increase in the capacity estimate for each condition. We report the first method here since a) it has one less free parameter and b) its interpretation is clearer.

Table 1. Subitizing capacities at each spacing in Experiment 1 (*Mean* ± *SEM*). Also shown are pairwise t-tests (Bonferroni corrected for multiple comparisons).

Spacing	5°	10°	20°	40°
5°	*1.73 ± 0.1*			
10°	$t(17) = 0.24$; $p = 1$	*1.77 ± 0.1*		
20°	$t(17) = 4.79$; $p < 0.001$	$t(17) = 5.62$; $p < 0.001$	*2.89 ± 0.2*	
40°	$t(17) = 7.89$; $p < 0.001$	$t(17) = 7.07$; $p < 0.001$	$t(17) = 3.31$; $p = 0.024$	*3.82 ± 0.3*

2.2. Results

To assess if crowding affects subitizing, we compared subitizing capacities and performance at each inter-object spacing (θ). Subitizing performance was high (mean ± SEM: 0.96 ± 0.01) at all spacings (see Supplementary Results S1.1 and Supplementary Fig. S1; $F(2.16, 36.7) = 0.38$, $p = 0.7$, ${}_p\eta^2 = 0.02$; Greenhouse-Geisser correction applied), suggesting that enumeration of small numbers of objects was equally and highly accurate at all distances. Importantly, we found that the subitizing capacities differed across spacing conditions ($F(2.27, 38.67) = 33.95$, $p < 0.001$, ${}_p\eta^2 = 0.67$). Subitizing capacities increased with spacing, as we would expect from crowding-like interactions between the objects. Follow-up pairwise comparisons (Table 1) indicated that there were no differences between capacities at the two closest (θ = 5° and 10°) distances. All other comparisons were significant. Interestingly, capacities were around 4 at the farthest spacing (θ = 40°), when the dots couldn't crowd each other, comparable to the well documented subitizing capacity in foveal vision (e.g., Trick & Pylyshyn, 1994). However, when the dots could crowd each other (θ = 5° and 10°) subitizing capacities were more than halved to around 1.7. That is, we observed considerable effects of internal crowding (Martelli, Majaj, & Pelli, 2005) on individuation, even when identification was not required. This is remarkable given that the dots were farther apart, even at the shortest spacing condition, than the resolving power of the visual system at that eccentricity. At 10 deg eccentricity, acuity (two-dot discrimination or letter identification) is about 0.15 - 0.3 deg (Anstis, 1974; Foster, Gravano, & Tomoszek, 1989), whereas our shortest spacing was 0.9 deg. In other words, even when the dots should have been distinguishable, the presence of items within Bouma's bound strongly impairs the visual system's ability to individuate them.

3. Experiment 2: External crowding of individuation

Experiment 1 showed that individuation could be internally crowded, conceptually replicating the impairment observed by Intriligator and Cavanagh (2001). A stronger test of whether individuation can be crowded in general would be to assess if irrelevant similar flankers impair subitizing. It is known that crowding is not only affected by the distance between targets and flankers (Bouma, 1970; Toet & Levi, 1992) but also by the similarity between them (e.g., Kooi, Toet, Tripathy, & Levi, 1994). Dissimilar flankers induce minimal or no interference with target identification. Here, we adopted the standard crowding paradigm to test if external flankers would also affect subitizing. The to-be-enumerated targets were surrounded by similar and dissimilar distracter flankers at various distances.

3.1. Method

In this experiment, an 18" CRT monitor with a resolution of 800 x 600 pixels and a refresh rate of 120 Hz was used. The distance to the monitor was fixed at 57 cm.

We planned to assess reaction times, rather than accuracies, so as to use the same measure as traditional subitizing studies. Thus, stimuli were presented until participants responded. To ensure that participants fixated well, we recruited four *experienced* observers.

We extended the range of tested numerosities to 9 here, because including more numerosities potentially allows for better bilinear fitting of data and hence for a more precise estimation of the subitizing capacity. It incidentally allows us to test the effect of flanker presence on counting (enumerating more than 4 objects; see Supplementary Results S2.3). Therefore, 1-9 target circles were presented within a 4x4 square grid centred at 10 deg in the lower visual field (Fig. 3). Each target circle had a diameter of 0.8 deg. The square grid was 6 deg on each side with the centres of adjacent cells 1.5 deg apart; hence adjacent targets would be within crowding distance of each other but well above two-dot discrimination thresholds. The specific locations of the target dots within the grid were randomly chosen on each trial. Flankers could be black squares (size 1 deg) or white X's (size 1 deg). The top two panels in Figure 3A illustrate these different flanker types. As can be observed in these panels, neither of these flankers could be mistaken for the targets: the black square flankers were larger and had a different shape relative to the black circle targets and the white flankers were dissimilar in shape, size and contrast polarity. The flankers were placed, one in each cardinal direction (left, right, top, bottom). They could appear at one of three distances from the centre of the target grid: 4.25 (near), 6 (intermediate), or 7.75 (far) deg from the centre of the target grid. At the tested eccentricity, only the near flankers were within Bouma's bound. We included two control conditions, as depicted in the bottom two panels of Figure 3A. In one, we presented a single large black square 'frame' centred at 10 deg eccentricity that enclosed the entire target grid. The sides of the black frame were located at the same distance as the flankers; hence the square frame could be of size 8.5, 12, or 15.5 deg on each side. The frame condition tested for the effect of the presence of extra black features in the stimulus. It was nevertheless not expected to crowd the dots, since crowding is sensitive to the similarity between targets and flankers (Kooi et al., 1994; Levi, 2008): dissimilar objects don't crowd each other. We also included a no-flankers condition, as another baseline.

Each numerosity, flanker-type and flanker-spacing condition combination was tested with 24 trials, resulting in a total of 2160 trials, spread over five sessions. Three of the sessions tested enumeration of dots flanked by either black (squares) or white (cross) flankers. The other two sessions tested enumeration in the presence of a square black frame or without flankers. Each session of 450 trials started with 19 practice trials (not included in data analysis). The order of condition combination within each session was randomised. Each session was broken up into 3 blocks.

The experienced observers were asked to strictly maintain fixation on a central cross and report the number of target dots as accurately and as quickly as possible. The display stayed on until response. The inter-trial-interval was 1 s.

If individuation is susceptible to crowding, then external flankers that are similar to the target dots should impair their subitizing, if they are within Bouma's bound of the targets. Hence, we would expect black flankers to reduce subitizing capacity, particularly at the nearest spacing. On the other hand, white flankers and the frame, being dissimilar to the target dots, should induce weak to no crowding and hence should minimally impair subitizing, at any distance.

3.2. Results

Accuracy in reporting the number of dots was reasonably high (> 0.85 proportion correct, pooled across all conditions and participants; see Supplementary Results S2.1 and Supplementary Fig. S2 for accuracy data). As planned, we analysed the reaction times for correct trials. We determined the subitizing capacity for each participant for each flanker-type and flanker-spacing condition combination (see Supplementary Fig. S3 for individual data). These were computed by fitting bilinear functions to individual mean correct reaction time data (Fig. 3B). Since we tested only four participants, we did not conduct the usual parametric tests on their data. We observe that, as can be seen in Figure 3B, reaction times were slower when target dots were surrounded by black flankers than in the other flanking conditions (white flankers, frame, or no flankers). This was particularly the case at the nearest flanking distance, which is within Bouma's bound. Interestingly, subitizing capacity (Fig. 3C) was most severely impaired when targets were surrounded by similar flankers at the closest spacing.

To test these observations, we subjected the reaction time data to a bootstrap analysis conducted separately for each participant (see Supplementary Results S2.2 for full details and Supplementary Fig 3B). In brief, on each of 1000 iterations, we sampled correct reaction times with replacement and estimated subitizing capacity for each flanker-type and spacing. We then determined the change in subitizing capacity as a function of spacing for each flanker-type. A slope of zero indicates that flanker spacing has no effect on subitizing, and hence that those flankers do not impair subitizing. Positive slopes indicate that flankers impair subitizing more at near distances, as expected from crowding. As depicted in Figure 3D, black flankers substantially affected subitizing (mean slope across all four participants: 0.35 ± 0.028 items/degree; *p*'s < 0.05 in 3 out of 4 participants), whereas the frame flanker did not (mean slope: 0.02 ± 0.036; all *p*'s>0.25). The effect of white flankers was marginal (*p*'s range from 0.07 to 0.18) and mild (mean slope: 0.13 ± 0.037).

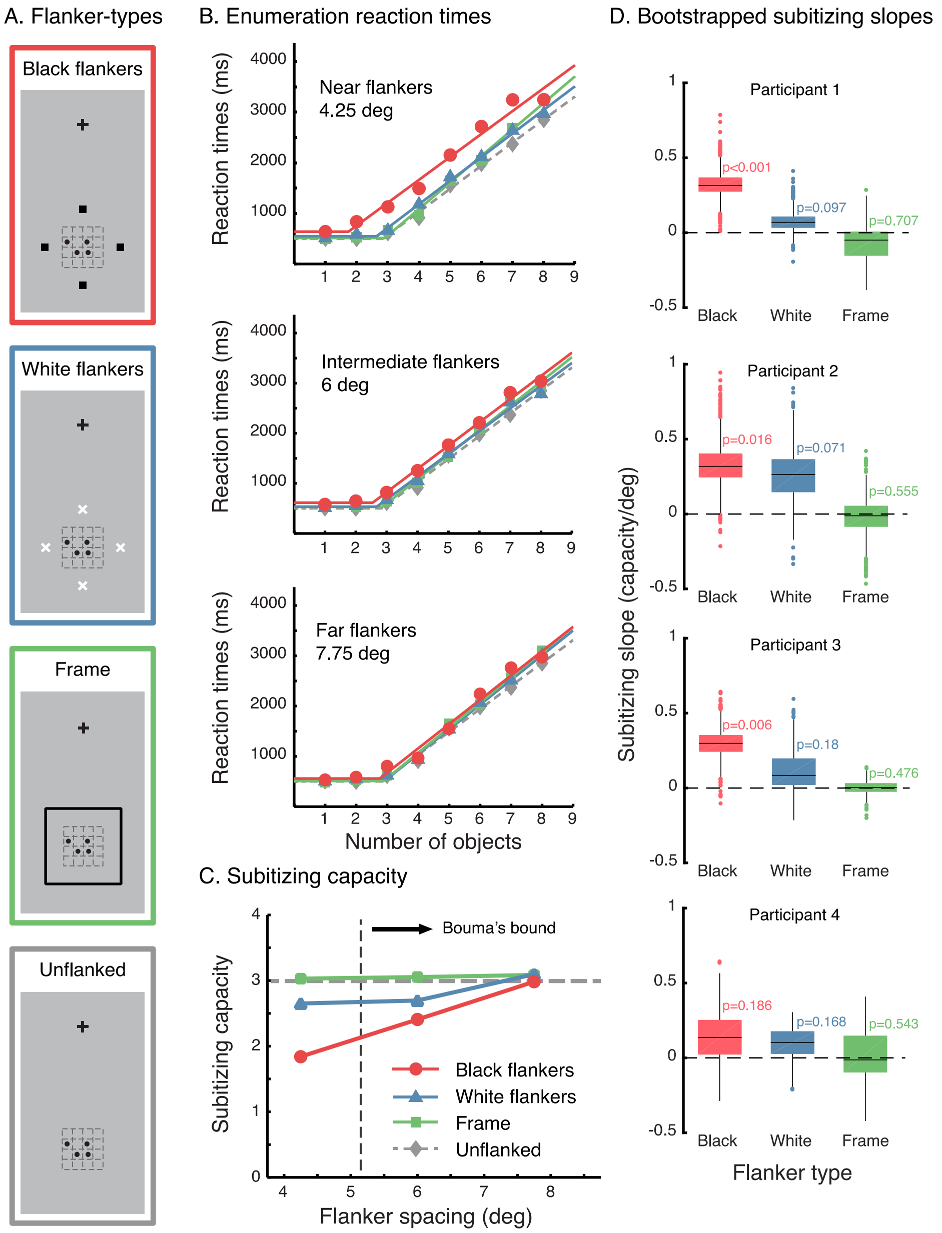

Figure 3: Stimuli and results for Experiment 2. A. 1-9 target circles were presented in a 4x4 grid centred at 10 deg eccentricity in the lower visual field (grid was not visible in the actual experiment). Three flanker-types (black flankers, white flankers, black frame) were tested at three different spacings. An unflanked condition was also included as a baseline (same data shown in all three plots in 3B; grey diamond markers and dashed line). **B.** Mean reaction times in reporting the number of items for each flanker-type at different flanker spacings. Bilinear fits are shown. **C.** Subitizing capacity for each flanker type as a function of flanker spacing. The horizontal dashed grey line indicates the subitizing capacity for the unflanked condition. Individual participant data is shown in Supplementary Fig. S3. **D.** Boxplots of bootstrapped slopes (change in subitizing capacity per deg of flanker spacing) for each flanker type for each participant. *p*-values adjacent to the corresponding boxplots indicate whether the bootstrapped distribution differs from a slope of zero.

In short, the closer the black flankers to the target dots, the larger the reduction in subitizing capacity. White flankers had a less dramatic effect, with only a mild influence of spacing, and the frame flanker had no effect at any spacing. In fact, subitizing capacity in the presence of white flankers and square frame were comparable to the subitizing capacity of unflanked targets, implying that mere presence of additional features did not affect subitizing. Subitizing capacity was substantially reduced specifically by similar flankers within Bouma's bound. These findings were further corroborated by complementary analyses conducted on subitizing performance (Supplementary Fig. S3, panel C), which estimates the reaction times at which subitizing occurs. We found that subitizing performance was the slowest within Bouma's bound, but only when the targets were surrounded by black flankers. These results provide evidence that crowding by irrelevant but similar flankers also impairs subitizing.

4. Experiment 3: Individuating complex objects

4.1. Experiment 3a: The effect of familiarity

Experiments 1 and 2 showed that subitizing is impaired when the target objects are close to each other or if they are surrounded by closely placed external flankers. However, the objects tested in the experiments were simple (squares and circles), similar to those typically used in enumeration tasks. These objects do not carry identifying information, whereas crowding is studied with identification tasks. If we desire to compare performance across enumeration and identification, we have to first determine if enumeration of more complex and identifiable objects can also be crowded. Hence, the current experiment was designed to test the generalisability of Experiment 2's findings by using more complex objects (letters). Further, we tested the effect of familiarity on the effect of crowding on subitizing by using upright and rotated letters. Upright letters are familiar and easier to identify, whereas rotating them makes them unfamiliar and harder to identify (Bergen & Julesz, 1983; Vanrullen, 2009).

4.1.1. Method

Thirteen observers participated in Experiment 3a. We used naïve participants instead of a small number of experienced observers used in the previous experiment. We also tested if the results of Experiment 2 can be replicated if we used accuracy measures instead of reaction times. We fixed the presentation duration to 160 ms and, as in Experiment 1, measured accuracy of reporting numerosity.

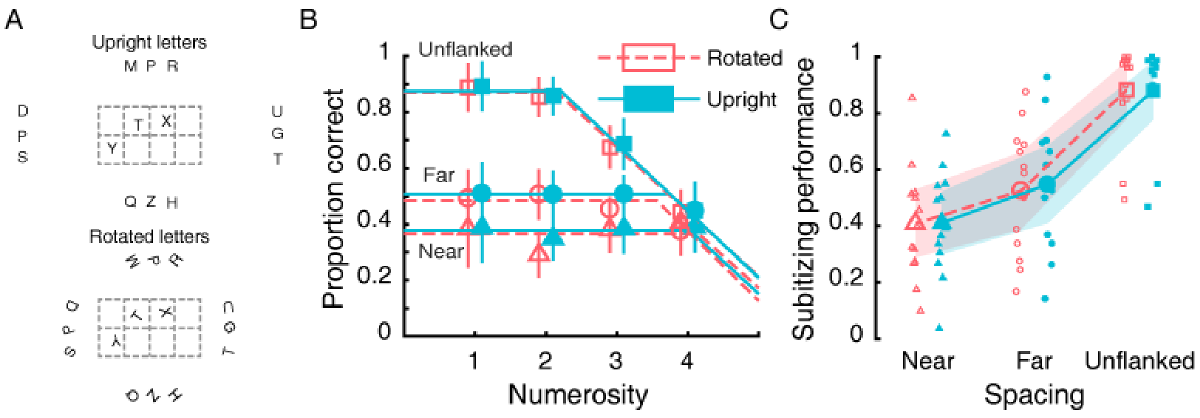

Figure 4: Stimuli and results for Experiment 3a. A. 1-5 target letters were presented within a 2x4 grid centred at 10 deg eccentricity either in the left or the right visual field (grid was not visible in the actual experiment). The letters were either in familiar (top panel) or unfamiliar (bottom panel) orientations. The flanker sets were either far (top panel) from or near (bottom panel) the targets. **B.** Accuracy of reporting the number of items presented for the two familiarity conditions at different flanker spacings. Bilinear fits are also shown. Error bars are 95% CI. **C.** Subitizing performance for the three spacing conditions with upright (red) and rotated (blue) letters data are jittered. Each dot represents data from one participant for that condition. The shaded areas are 95% CI.

The distance to the monitor was fixed at 57 cm. Targets and flankers were black letters on a white (91.5 cd/m^2) background (Fig. 3). 1-5 letters were presented within a 2x4 grid, centred at 10 deg eccentricity. Each cell in the grid was 1.2 deg on each side. Letters were 0.67 deg in size and were randomly allocated to these cells. There were three flanker conditions: near, far, or no flankers. When presented, flankers appeared in four separate sets, one in each cardinal direction from the target grid. Each flanker set was a cluster of three letters: arranged vertically (grid size 3.6 by 1.2 deg) in the left and right positions and horizontally (size 1.2 by 3.6 deg) in the top or bottom positions. Each cell in these sets was filled by a randomly chosen letter. In the far flanker condition, the left and right distracter sets were placed 5 deg from the centre of the target grid, and the top and bottom distracter sets were 2.5 deg from the centre of the target grid. Thus, they were at the edge of Bouma’s bound. In the near flanker condition, all four sets were 2.5 deg from the centre of the target grid and were hence within Bouma’s bound. There were two familiarity conditions: upright letters and rotated letters. In the upright letters condition, the targets and flankers were in the familiar upright orientation. In the rotated letters condition, each letter was randomly rotated between 45^o and 315^o of vertical, ensuring that they were always seen in a non-familiar orientation. Each condition combination was tested with 40 trials (total 1200 trials).

The procedure was the same as in Experiment 1 and participants were once again asked to report the number of target letters. The only differences compared to the procedure used in Experiment 1 were that 1) the targets and flankers were presented either in the left or the right hemifield, randomly chosen and 2) no time limit for responding was imposed.

4.1.2. Results

We examined the effect of external flankers on subitizing by applying bilinear fits to the accuracy data as a function of numerosity, for each spacing and familiarity (letter rotation: upright or random) condition. For subitizing capacities to be a valid measure, subitizing performance (the highest performance that participants were capable of at low numerosities) should remain high (say, > 0.9 proportion correct). However, it was evident that subitizing performance was substantially less than 0.9 in some spacing conditions. Thus, we did not analyse subitizing capacities and focused only on subitizing performance. We found that target-flanker spacing substantially affected subitizing performance ($F(2, 24) = 40.28$, $p < 0.001$, ${}_p\eta^2 = 0.77$). This seems to be mainly driven by high performance in the absence of flankers (0.88 ± 0.05) in contrast to considerably reduced performance in the presence of flankers (far flankers = 0.54 ± 0.06, $t(12) = 6.12$, $p < 0.001$; near flankers = 0.41 ± 0.05; $t(12) = 7.63$, $p < 0.001$). Near flankers reduced performance even further than far flankers ($t(12) = 2.92$, $p = 0.039$). These results indicate that subitizing of complex objects is considerably impaired and hence crowded by the presence of flankers.

Familiarity had no effect on subitizing performance ($F(1, 12) = 1.7$, $p = 0.22$, ${}_p\eta^2 = 0.12$). Further, there was no interaction between spacing and familiarity ($F(2, 24) = 0.72$, $p = 0.5$, ${}_p\eta^2 = 0.06$). These results suggest that target-flanker spacing affects subitizing, as one would expect from crowding-like effects on subitizing, but this interference was not modulated by familiarity. The latter finding indicates that subitizing and its impairment takes place before familiarity of an object is ascertained and hence likely before feature integration.

The effect of complexity on subitizing can be determined by comparing subitizing capacities across simple and complex objects under similar testing conditions. Experiment 1 tested enumeration of briefly presented dots and the current experiment tested briefly presented upright and rotated letters. The unflanked data in the latter (inter-letter spacing 1.2 deg) are directly comparable to closely spaced dots in the former (inter-dot spacing 0.9 deg or 1.7 deg). Subitizing capacity for dots was ~1.7 - 1.8 items at these spacings. Capacity, determined from the bilinear fits, was 2.3 ± 0.2 items for upright letters and 2.1 ± 0.1 items for rotated letters. These capacities are comparable (all *p*'s > 0.4) indicating that subitizing in the periphery appears to be independent of the type, complexity and familiarity of the objects being enumerated.

4.2. Experiment 3b: Does the task assess individuation?

An alternative interpretation of the finding that subitizing is affected by target-flanker distance might be that the impairment is not due to crowding or not just attributable to crowding, but due to inadvertently including one or more flankers in the enumeration process, at least on some of the trials. This can occur because the targets and flankers are physically indistinguishable, except for their position. Additionally, there is no physical boundary that helps participants know which objects are targets to be enumerated and which ones are flankers to be ignored. Wender and Rothkegel (2000) argued that segmenting the target objects from the distractors is an essential step for enumeration in the subitizing range. When this segmentation fails, for example, when the targets and flankers are spatially interspersed, subitizing is severely impaired. Although the targets and flankers were spatially separated in our experiments, the objects were presented in the periphery where spatial localisation is poor. Hence, target-flanker segmentation could have failed, particularly in the presence of close similar flankers, leading to incorrect enumeration. In other words, it might be possible that the participants mistook some of the flankers for the targets, or the targets for the flankers, at least some of the time. This would lead to errors in enumeration, compared to when target-flanker segregation is straightforward. Further, this confusion would be higher at shorter target-flanker distances, explaining our results. Note that the inclusion of flankers or exclusion of targets can itself be argued to be crowding due to a failure of selective attention (He et al., 1996). With that caveat in mind, it would nevertheless be useful to determine if this strategy is being (mal-)utilised by the visual system. Another possibility is that the inadvertent and automatic inclusion of one or more flankers in the enumeration process might trigger the operation of the approximate number system, rather than the more precise subitizing process leading to further errors (Feigenson, Dehaene, & Spelke, 2004).

We tested this possibility in the current experiment by making two changes to Experiment 3a. We used only the letter X as flankers and the targets were randomly selected from the entire alphabet except X. Thus, the targets and flankers should be distinguishable, making their segmentation easier (Goldfarb & Levy, 2013; Mazza, Turatto, & Caramazza, 2009). Participants were explicitly instructed about the differences in the identities of targets and flankers. Further, in some blocks, we introduced a green circular frame that delineated the location of the targets. All targets were enclosed within this circle and flankers would always be outside it. Participants were asked to enumerate only those objects within this large 'ring'. These two factors should enhance the ability to segregate the targets from flankers and hence mitigate the effects of confusing the targets and flankers for each other.

4.2.1. Method

Nineteen naïve observers participated in Experiment 3b. Data from one participant was discarded as their performance was low for all numerosities in all conditions. The material and stimuli were the same as in Experiment 3 except for the following changes. Flankers and targets were upright letters; we did not use rotated letters. Flankers were always Xs, and targets could be any letter in the alphabet except X. In some blocks two green rings of diameter 3.5 deg were presented, one on each side of fixation, centred at 10 deg eccentricity. These rings, when present, encircled the targets but excluded the flankers, thus serving as cues to segregate targets and flankers. Note that we do not expect any crowding of the target letter(s) from these rings for a few reasons. 1) Experiment 2 showed that a black frame did not interfere with the targets, 2) the rings have a different colour and shape from the targets. It has been shown that if flankers and targets differ on some feature dimension, crowding is weaker (Kennedy & Whitaker, 2010; Kooi et al., 1994), 3) the rings are present throughout the block, without break. It is known that previewing distracters reduces or eliminates crowding (Greenwood, Sayim, & Cavanagh, 2014; Scolari, Kohnen, Barton, & Awh, 2007). There were a few other differences, relative to Experiment 3. First, we used 6 numerosities. Second, the stimuli were presented for 150 ms. Third, the monitor was placed 50 cm from the observer.

In this experiment, we tested six conditions: each condition was a combination of a) 2 ring presence options (targets enclosed by a ring or not) and b) 3 flanker spacing (2.5 deg, 5 deg or unflanked). Each condition and numerosity combination was tested with 36 trials (total of 1296 trials).

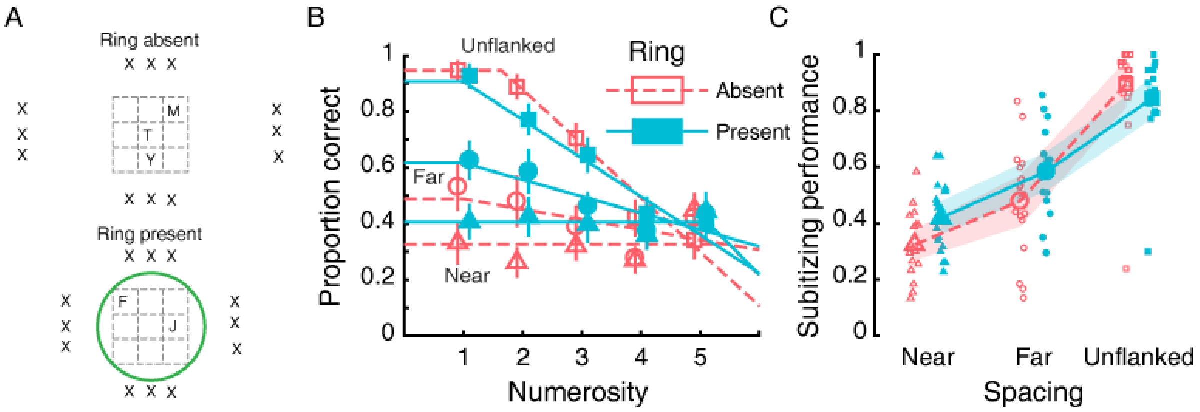

Figure 5: Stimuli and results for Experiment 3b. A. 1-6 target letters were presented within a 3x3 grid centred at 10 deg eccentricity either in the left or the right visual field (grid was not visible in the actual experiment). The target letters were either not enclosed (top panel) or enclosed (bottom panel) within a green ring. The flanker sets were all triplets of Xs, either far (top panel) from or near (bottom panel) the targets. **B.** Accuracy of reporting the number of items presented for each ring presence condition at different flanker spacings. Bilinear fits are also shown. Error bars are 95% CI. **C.** Subitizing performance for the three spacing conditions: ring absent (red) and present (blue) letters data are jittered. Each dot represents data from one participant for that condition. The shaded areas are 95% CI.

The procedure was the same as in Experiment 3a except for the following differences. In blocks where rings were presented, the two rings, one on each side of fixation, were present throughout the block. The order of blocks (with and without rings) was randomised.

4.2.2. Results

Once again, we found that target-flanker spacing strongly affected subitizing performance ($F(2, 34) = 146.6$, $p < 0.001$, ${}_p\eta^2 = 0.9$). Subitizing was high in the absence of flankers (0.9 ± 0.01) in contrast to considerably reduced performance in the presence of flankers (far flankers = 0.55 ± 0.03, $t(17) = 10.51$, $p < 0.001$; near flankers = 0.38 ± 0.02; $t(17) = 17.62$, $p < 0.001$). Near flankers reduced performance even further than far flankers ($t(17) = 6$, $p < 0.001$). This effect of flanker distance was observed even when we restrict the results to the ring-present conditions, where we expect improved target-flanker segregation and reduced possible confusions (no flankers = 0.87 ± 0.02; far flankers = 0.6 ± 0.03; near flankers = 0.43 ± 0.03). These results indicate that subitizing is impaired by the presence of flankers.

The presence of a ring influenced subitizing performance ($F(1, 17) = 7.03$, $p = 0.017$, ${}_p\eta^2 = 0.29$), where the ring modestly improved subitizing performance (ring present, 0.64 ± 0.02; ring absent, 0.59 ± 0.02). However, there was a significant interaction between spacing and ring presence ($F(2, 34) = 20.69$, $p < 0.001$, ${}_p\eta^2 = 0.55$). To determine the source of this interaction, we conducted pairwise tests between subitizing performance with and without rings at each flanker spacing separately. Subitizing performance for unflanked letters was higher in the absence of a ring (0.93 ± 0.02) compared to when a ring was present (0.87 ± 0.02; $t(17) = 3.36$, $p = 0.002$), indicating that the ring impaired subitizing to a small extent. In contrast, subitizing performance was higher in the presence of a ring for both far (ring present: 0.6 ± 0.03, ring absent: 0.5 ± 0.04; $t(17) = 3.2$, $p = 0.005$) and near (ring present: 0.43 ± 0.03, ring absent: 0.33 ± 0.03; $t(17) = 4.47$, $p = 0.002$) flankers. That is, the ring enhanced subitizing in the presence of flankers, perhaps by aiding the segregation of target and flanker letters. These findings suggest that some of the effects of flankers on subitizing, observed here and in the previous experiments, can be explained by the inability to distinguish flankers from targets. However, the influence of this confusion appears to have been mild to moderate, with an improvement in performance of only a few percentage points (~10% on average), relative to the effect of flanker spacing.

The current experiment, using two cues to augment segregation between targets and flankers, nevertheless found that flanker distance substantially affected subitizing, indicating that subitizing can be crowded. The impairment in performance is not merely due to the inability of the visual system to distinguish flankers from targets. The experiment does not completely rule out the possibility that confusions might have persisted despite the two segregation aids, but is strongly indicative that confusions might not be the driving force in the effect of flanker distance on crowding and that *subitizing* is impaired by flanker presence.

6. Experiment 4: Comparing subitizing and identification

In Experiment 4, we tested both subitizing and identification in the presence of flankers using a comparable stimulus setup. This allows us to closely probe the relationship between the two phenomena. Here, we determined the critical spacing for target identification and enumeration in the same participants. Critical spacing is a commonly used measure of crowding and is typically defined as the minimal distance between targets and flankers required to achieve a threshold level of performance. If crowding occurs at the individuation stage, as the attentional hypothesis predicts, then impairment of identification and subitizing by flankers should be comparable; that is, their critical spacings should be the same or nearly the same. However, given the substantial differences in task requirements, it is possible that other factors (e.g., visual short-term memory, task difficulty) might supervene upon the behavioural outcomes. Hence, we expect the critical spacing for the two tasks to at least correlate. Thus, a strong test of our hypothesis would be the that the critical spacing in the two tasks would be the same or proportional (slope = 1), whereas a weaker test would be a positive correlation with a slope less than 1. We should note that a correlation would indicate a potential common mechanism, even if it does not imply it.

6.1. Method

6.6.1. Participants

Since we wanted to compare critical spacing across tasks, we used a larger sample size with twenty-four observers participating in this experiment.

6.1.2. Material and stimuli

The materials were the same as in Experiment 3b with a few changes. Participants completed two tasks, identification to assess crowding, and enumeration to assess subitizing. In the identification task, a single letter chosen from K, N and V was used as a target. This letter was presented at 10 deg eccentricity. In the enumeration task, 1-3 of these letters were presented, with replacement. These target letters were presented in randomly chosen cells of a 3x3 grid. Flankers were triplets of Xs, as in the previous experiment, presented at one of six distances: 2.25, 2.5, 3, 3.5, 4.5 and 6 degrees. An unflanked condition was also included. As in experiment 3b, two green rings were presented in half the blocks.

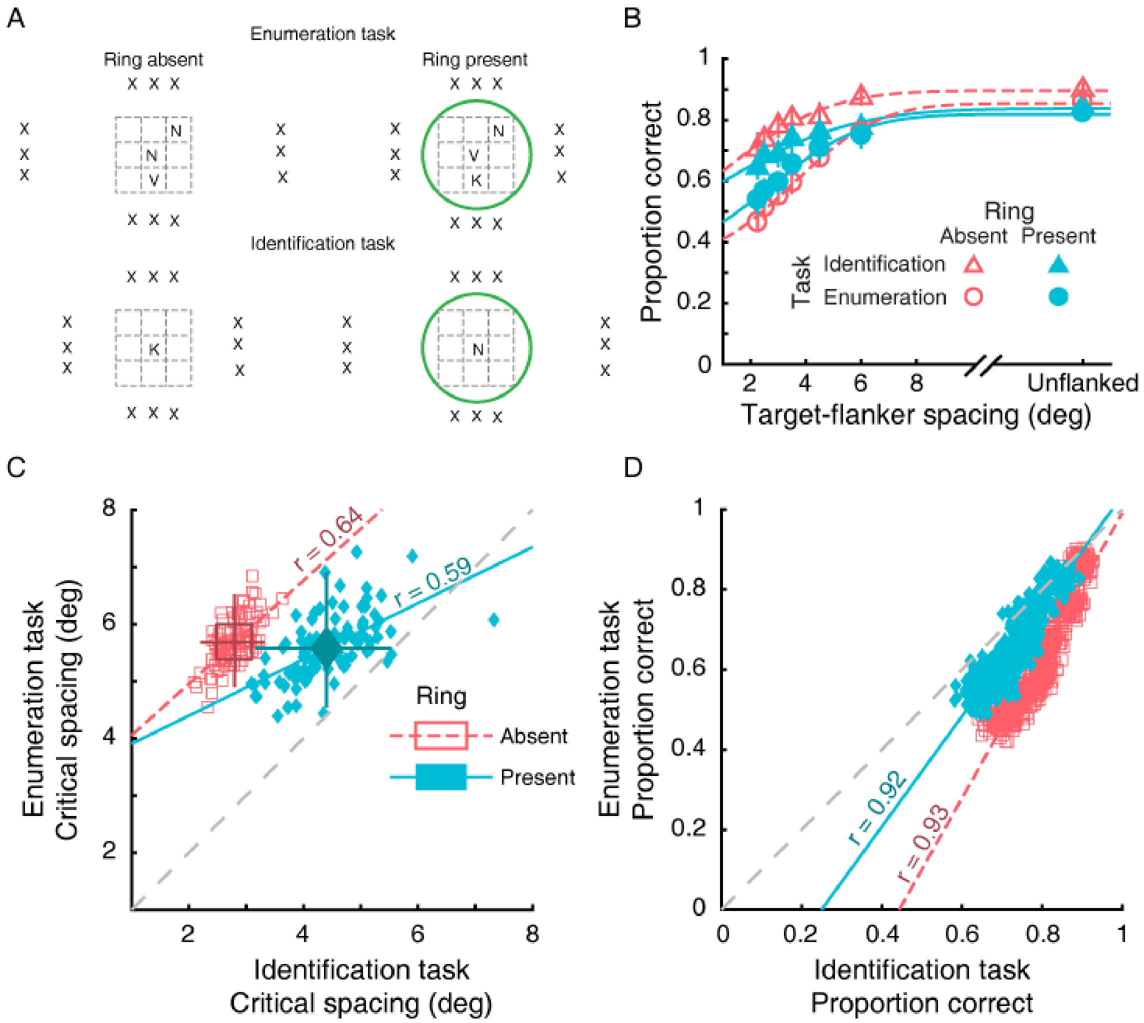

Figure 6: Stimuli and results for Experiment 4. A. Participants were asked to report the number of target objects in some blocks (top row in panel A) and to identify the solitary target presented in others (bottom row). In the enumeration task, 1-3 target letters were presented within a 3x3 grid centred at 10 deg eccentricity either in the left or the right visual field (grid was not visible in the actual experiment). In the identification task, one letter was presented in the centre of the grid. In both cases, the target(s) were selected from the letters K, N, and V. In some blocks, a green circle separating the target grid from the flankers (right column) was presented. **B.** Accuracy of reporting the number or identity of the target(s) as a function of target-flanker spacing. Cumulative Gaussian fits are shown. Error bars are 95% CI. **C.** Scatterplot of critical spacing in the two tasks estimated from bootstrapped psychometric curves. Red squares represent critical spacing estimates when no ring was presented, and blue diamonds are critical spacing estimates when a ring was presented. The mean and 95% CI are shown as larger and darker symbols with error bars. Linear fits for the bootstrapped estimates are also shown along with correlation coefficients. The dashed grey line represents the equality line, with a slope of 1. **D.** Scatter plot of proportion correct values in the two tasks at all target-flanker spacings. Best fitting straight lines and corresponding correlation coefficients are shown.

6.1.3. Procedure

Identification and enumeration tasks were tested in separate blocks. Half of each of these blocks included a green ring that separated the target(s) from the flankers. The order of these blocks was randomised. In each task, the target and flankers were presented for 150 ms either in the left or the right hemifield. Participants were then presented with 3 onscreen response options: K, N, and V in the identification task, or 1, 2, and 3 in the enumeration task. They were asked to select one of these options with the mouse. Each condition combination (task x ring presence x flanker spacing) was tested with 40 trials (total 1120 trials).

6.1.4. Data analysis: Psychometric curve fitting

We used a bootstrap approach to determine whether critical spacing in the two tasks were correlated. To do so, we sampled, with replacement, twenty-four participants and averaged their performance. We then fit cumulative Gaussians (Eq. 1) to the averaged accuracy data as a function of target-flanker spacing for each of the four conditions (2 tasks x 2 ring presence conditions).

$$y = \gamma + \left(\frac{(\alpha - \gamma)}{2} erfc \left(\frac{-\sigma(x - \mu)}{\sqrt{2}} \right) \right) \qquad (1)$$

where y is accuracy, x is spacing, γ is chance performance (0.33 here), α is the asymptote, σ is the slope of the psychometric curve, μ is the midpoint of the curve, and erfc is the complementary error function.

From these fits, we extracted critical spacing for each condition, that is, the spacing at which accuracy was at a certain level (0.75; using other values also gave similar results). We repeated this procedure 100 times. We then correlated critical spacing between enumeration and identification tasks for each ring presence condition separately. Please see Supplementary Results S3 for the non-bootstrapped procedure, showing the same results.

6.2. Results

Critical spacing was higher (crowding occurred over a larger distance) for the subitizing task than for the identification task for both ring present (subitizing = 5.58 deg ± 0.56; identification = 4.4 ± 0.67) and absent (subitizing = 5.68 ± 0.41; identification = 2.8 ± 0.29) conditions. This could reflect differences in difficulty between tasks. In other words, it was easier to identify a letter than to enumerate 1 - 3 letters. This is interesting as subitizing is often considered to be earlier and require less processing than identification. It might partially be explained by internal crowding (Experiment 2, current study; Martelli et al., 2005) between the multiple targets in the enumeration task. The enumeration has more than one object in two-thirds of the trials, increasing the chances of internal (target-target) crowding, leading to more errors.

Importantly, to assess the relationship between the two tasks, we determined the correlation between the critical spacings in these tasks. We found a strong correlation irrespective of whether a ring was present ($r = 0.59$, $p < 0.001$, slope = 0.49) or absent ($r = 0.64$, $p < 0.001$, slope = 0.9). The correlation is slightly higher when a ring is absent than when it is present, but the slope is much steeper and closer to proportionality when a ring is absent than when it is present. Looking closely at this change in slope, it is clear that the presence of the ring increased the critical spacing in the identification task but did not alter critical spacing in the enumeration task. The latter - no effect of ring on enumeration - replicates the findings of Experiment 3b. The increase in critical spacing (stronger crowding) in the identification task in the presence of a ring is puzzling, as the shape and colour of ring were specifically chosen to avoid crowding the target. It is even more puzzling since the ring was expected to reduce target-flanker confusions and hence reduce crowding that might be attributable to flanker substitution. The best explanation for this anomalous finding might be super-crowding of the target (Vickery, Shim, Chakravarthi, Jiang, & Luedeman, 2009). Vickery et al. found that when a target and its flankers were separated by a small ring or frame, critical spacing far exceeded Bouma's bound. They argued that this was because the ring-induced mild masking of the target extended the zone of interference between the target and its flankers. A similar interaction might have occurred in our display, increasing the critical spacing from ~2.8 deg without the ring to ~4.4 deg with a ring, a 50% increase. However, we must note that, unlike in the Vickery et al. experiments, our ring stayed on the screen throughout the block and hence any sort of lateral or backward masking should have been minimal. Nevertheless, if we assume that the solitary target was mildly masked, the binding between its features might be loosened more than otherwise and rendered the target susceptible to flanker interference. On the other hand, this putative masking by the flankers might not be sufficient to prevent the individuation of multiple letters, as tight binding is not necessary for this process. Hence enumeration would be preserved. This interpretation is supported by the finding that the ring reduced identification (ring absent = 0.9 ± 0.01, ring present = 0.84 ± 0.02; $t(23) = 3.41$, $p = 0.004$) but not enumeration (ring absent = 0.86 ± 0.02, ring present = 0.83 ± 0.01; $t(23) = 1.83$, $p = 0.16$) accuracy even in the absence of flankers. Thus, in this experiment, if we were to consider the ring absent condition to be less contaminated by other processes, such as super-crowding, then it is clear that the critical spacing is the same (slope ~1) in both tasks. This supports the hypothesis that crowding occurs at the individuation stage. Even in the ring present condition, where the slope is not 1, the correlation is high, and this passes the weaker test of the hypothesis.

We also conducted an alternative correlation analysis where we directly compared accuracy in the two tasks irrespective of target-flanker spacing. We found, once again, strong correlations between performance in the two tasks (with ring: $r = 0.92$, slope = 1.38; without ring: $r = 0.93$, slope = 1.78). That is, when accuracy was low in one task (probably due to the presence of close flankers), accuracy was low in the other task as well. Slopes were greater than 1 indicating that performance in the subitizing task fell off faster than in the identification task, complementing the critical spacing findings above.

7. General Discussion

Object recognition and individuation have typically been studied separately, with very few exceptions (e.g., Xu & Chun, 2009). In this study, we brought together these two distinct domains to illuminate the processing pipeline for object recognition. We tested if subitizing, an individuation process, could be crowded, a process where recognition fails. We found that subitizing of objects is impaired when they are close to each other (Experiment 1), conceptually replicating Intriligator and Cavanagh’s (2001) findings. Crucially, subitizing is also impaired if the target objects are surrounded by irrelevant but perceptually similar flankers (Experiments 2-4). This impairment seems to be the same for objects of differing familiarity (Experiment 3). Further, the flanker induced interference was comparable for both subitizing and crowding tasks (Experiment 4), suggesting that individuation and identification share a common processing pathway.

These findings recommend that an individuation stage should be included in the object recognition pathway (Fig. 1B) and that crowding occurs by impairing individuation. This does not mean that additional interference at the feature integration stage or later does not occur (cf. Louie, Bressler, & Whitney, 2007; Manassi & Whitney, 2018). But the parsimonious explanation is that interference occurs primarily at the individuation stage. Remarkably, in EEG studies, the neural signatures of both crowding and individuation are observed in the same electrodes (occipital) at around the same time, around 200 ms after stimulus onset (crowding: Chicherov, Plomp, & Herzog, 2014; individuation: Mazza, Pagano, & Caramazza, 2013). This further supports the notion that crowding is due to interference at the individuation stage.

7.1. Crowding and selective attention

Our results also lend support to the proposal that crowding is a consequence of attentional limitations. Only the attentional hypothesis of crowding (He et al., 1996; Intriligator & Cavanagh, 2001) explicitly predicts that individuation can be crowded, which is what we find. Further, there is accumulating evidence that subitizing can be attributed to limitations in attention (Egeth, Leonard, & Palomares, 2008; Mazza & Caramazza, 2015; Vetter et al., 2008). That is, the same mechanism, attentional limitation, can explain crowding and subitizing. Selective attention is resource limited and can typically select 3-4 objects – subitizing. However, each act of selection has a resolution limit. If multiple objects are closely located, they cannot be separately resolved – crowding.

The attentional theory might have an additional advantage in explaining recent findings that seem to challenge pooling-based accounts. Studies from Herzog’s group (Herzog & Manassi, 2015; Herzog et al., 2015) and elsewhere (e.g., Levi & Carney, 2009; Livne & Sagi, 2007; Poder, 2006) have shown that a) increasing the number of objects can sometimes reduce crowding, b) grouping among flanking objects can reduce crowding, and c) flankers far beyond Bouma’s bound can influence target identification. Standard bottom-up pooling models cannot account for these findings. Interestingly, participants’ performance appears to be well correlated with subjective ratings of the *appearance* or Prägnanz of the stimulus array. Hence, it has been argued that top-down signals based on perceptual grouping allow the target to be segmented away from the grouped flankers, thus reducing crowding in the conditions noted above (Francis et al., 2017). This account is consistent with the ability of attention to individuate objects based on perceptual grouping (Roelfsema & Houtkamp, 2011), which in turn would alleviate crowding. If the target is separated by attention from the grouped flankers, then the interference between them would be minimised.

How can attention segment targets from flankers through grouping? Roelfsema and Houtkamp (2011) argue that grouping occurs by two processes, which they term *base* and *incremental* grouping, respectively. If there are neurons in the early visual system specialised for certain features or feature conjunctions (say a red horizontal line), then elements with these characteristics are automatically and rapidly grouped together, in parallel across the entire visual field. This form of grouping often occurs according to Gestalt principles (Koffka, 1935) and is called *base* grouping. Some of the known effects of grouping on crowding, such as the effect of similarity between targets and flankers or that of configural grouping of flankers (Kooi et al., 1994; Livne & Sagi, 2007) might be attributed to base grouping. Here, flankers automatically group with each other, allowing attention to select the target unimpaired.

On the other hand, when there are no neurons specialised to process the stimuli, grouping can still occur between such elements, but with effortful, time-consuming allocation of attention. This slow attention-based grouping is termed *incremental* grouping. Attention, here, operates through top-down selection signals guiding the grouping process by incrementally shifting the focus of selective attention along an object’s contour or surface to segment it from its background. Neurons in the early visual cortex might not be specialised to process several of the stimuli used in the Herzog lab (e.g., regularly spaced complex but similar shapes). Given that the presence of such objects seems to violate the expectations from standard pooling models, it can be argued that their effect manifests through incremental grouping. This incremental grouping also leads to appearance change. Our finding that individuation is necessary for object recognition and is likely to require attention lends support to this interpretation.

7.2. Alternative explanations

7.2.1. Crowding due to feature pooling

Does the above analysis rule out non-attentional theories of crowding? We argue that the data strongly support the attentional hypothesis. However, some non-attentional theories of crowding, such as pooling or averaging, might, with modifications, be able to accommodate our findings. The comparability of impairment in both subitizing and identification tasks indicates that individuation and identification must share a common pathway. This places constraints on the modifications that can be applied to the two-stage pooling model in order to account for our results.

First, as discussed in the introduction, pooling can push the crowded objects closer together (Korte, 1923). It is possible that this leads to underestimation, because the perceived locations might be too close for the visual system to resolve. Although our experiments cannot categorically exclude this possibility, we do not think that this is the case, since 1) the visual system's two-dot resolution is much finer than the distances tested in crowding, and 2) the compression observed in crowding studies do not seem to lead observers to report fewer objects (Sayim & Wagemans, 2017). Hence it is unlikely that spatial compression due to pooling can explain our results.

Second, the pooling hypothesis could be potentially modified to account for our data. If an individuation stage were introduced prior to the feature integration stage, the presence of flankers might prevent or interfere with this individuation process through a bottom-up process. Hence, a non-attentional pooling mechanism might also explain our results. Such a modified proposal closely resembles the attentional hypothesis at this point. It must be noted, however, that it is unclear how feature pooling or averaging can lead to interference at the individuation stage. That is, a clear bottom-up mechanism for interference at the individuation stage will have to be developed before this modified proposal can be considered viable.

7.2.2. Subitizing by pattern recognition

Another alternative explanation of our finding that subitizing is susceptible to crowding is that enumeration in the subitizing range is not due to individuation but relies on a sort of pattern or shape recognition (Krajcsi et al., 2013; Logan & Zbrodoff, 2003; Mandler & Shebo, 1982). According to this explanation, small numbers are enumerated rapidly because they form the vertices of simple and familiar shapes such as triangles and quadrilaterals. Therefore, subitizing can be crowded since flankers can readily impair such (virtual) shape recognition. That is, flanker interference occurs at the feature integration stage and does not require an individuation stage.

This view of subitizing as pattern recognition is not widely supported. The pattern recognition hypothesis relies on extracting numerosity from the initially perceived virtual shape. Hence, as Trick and Pylyshyn (1994) noted, enumeration should be impaired when, say, three objects are collinear. Since the pattern in this case is linear, it should elicit the value ‘two’. However, participants are fast and accurate while enumerating three collinear objects (Trick, 1987). Further, the chief evidence supporting the pattern recognition hypothesis (Krajcsi et al., 2013; Mandler & Shebo, 1982; Wender & Rothkegel, 2000) is that dots presented in a ‘canonical’ pattern (akin to dice) are easier (faster and more accurate) to enumerate than when they are randomly organised. There are several differences between these two types of stimuli that might have led to these results, such as familiarity, overlearning, symmetry, perceiving or rapidly learning to perceive highly organised shapes as symbols representing specific numbers. Hence, they might not be engaging a true enumeration process. In short, the evidence and logic of the pattern recognition hypothesis of subitizing is not compelling.

On the other hand, there is evidence that subitizing requires individuation of each object (see Mazza & Caramazza, 2015 for a review). Recent studies (Ester, Drew, Klee, Vogel, & Awh, 2012; Mazza & Caramazza, 2015; Mazza et al., 2013) have shown that a lateralized EEG component called N2pc, which is known to index attentional selection, monotonically increases with numerosity up to 4 items. This has been taken to suggest that subitizing requires object individuation. N2pc also increases as the tracking capacity increases in multiple object tracking (MOT; Drew & Vogel, 2008), suggesting that the same individuation process occurs in subitizing and MOT. However, the virtual shape connecting objects distorts over time in MOT and would be hard to maintain (Yantis, 1992), arguing against a pattern recognition model of individuation. Further, the subitizing span is closely tied to visual short-term memory capacity (Knops, Piazza, Sengupta, Eger, & Melcher, 2014; Piazza, Fumarola, Chinello, & Melcher, 2011). This is inconsistent with the possibility that only one virtual shape is being recognized, but is consistent with the hypothesis that multiple items are individuated in subitizing. Additionally, Anobile et al. (2015) reported that objects need to be perceptually segregate-able (particularly at low numerosities) for good enumeration performance, again suggesting that individuation is necessary for enumeration. Interestingly, in Experiment 1 of our study, and also in Palomares et al.'s (2011) study, items were arranged along an imaginary circle, which would have made it difficult to form a virtual shape. Yet, subitizing was robust at large inter-item distances.

More importantly, if subitizing is subserved by the recognition of a single large virtual shape, then this large shape should be crowded by a similar large shape and not by small dissimilar shapes (Harrison & Bex, 2015; Kooi et al., 1994; Pelli et al., 2004). Contrary to this expectation, in Experiment 2, a large black frame flanker barely made any difference to enumeration performance, whereas nearby small black square flankers, which are quite dissimilar to the large virtual shape, substantially impaired subitizing. These findings and arguments, taken together, suggest that subitizing as tested in our study is not based on pattern recognition but on the individuation or selection of separate objects.

7.2.3. Is subitizing impaired?

A third possible explanation is that it was not subitizing that was impaired by flankers. The impairment might instead be attributed to participants' inability to effectively segregate flankers from targets because they are similar and close to each other. If the targets and flankers are not distinguishable, participants might consider some of the flankers to be targets, or vice versa, leading to enumeration errors. This confusion increases with decreasing target-flanker distance, explaining the current results. We tested this possibility, in Experiments 3 and 4, by attempting to minimise or eliminate the target-flanker confusions. We did so by using physically different letters as flankers and distracters (and informing the participants about it) and also by separating them with a boundary. Flanker induced impairment of enumeration nevertheless remained high, allowing us to rule out target-flanker confusion as the chief or only source of the impairment. There was a small improvement in performance indicating that such confusion did play a small role in the impairment observed in the previous experiments (1-3a). These findings argue that participants, for the most part, did not confuse the targets and flankers. Their visual system was individuating and subitizing the target objects, which was impaired by the presence of flankers.

8. Conclusion

In this study, we found evidence that crowding, a recognition specific phenomenon, substantially impaired subitizing, an individuation specific phenomenon. These results suggest that a) individuation is an essential stage in the object recognition pipeline, and b) further supports the proposal that both crowding and subitizing are due to limitations of selective attention.

Acknowledgements

We thank Marlene Poncet and Patrick Cavanagh for helpful suggestions on the manuscript. We thank Louise Hill for help with some of the data collection. We are grateful to the reviewers and editor for their insightful comments on the framing of the study and our interpretation of the results, which led us to rethink the implications of our findings.

References

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision*, *15*(5), 4. https://doi.org/10.1167/15.5.4

Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14*(7), 589–592.

Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing: Magical numbers or mere superstition? *Psychological Research*, *54*(2), 80–90.

Bergen, J. R., & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man, and Cybernetics*, (5), 857–863.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, *94*(2), 115.

Bouma, H. (1970). Interaction Effects in Parafoveal Letter Recognition. *Nature*, *226*, 177–178. https://doi.org/10.1038/226177a0

Chakravarthi, R., & Cavanagh, P. (2009). Recovery of a crowded object by masking the flankers: Determining the locus of feature integration. *Journal of Vision*, *9*(10), 4–4. https://doi.org/10.1167/9.10.4

Chaney, W., Fischer, J., & Whitney, D. (2014). The hierarchical sparse selection model of visual crowding. *Frontiers in Integrative Neuroscience*, *8*. https://doi.org/10.3389/fnint.2014.00073

Chicherov, V., Plomp, G., & Herzog, M. H. (2014). Neural correlates of visual crowding. *NeuroImage*, *93*, 23–31. https://doi.org/10.1016/j.neuroimage.2014.02.021

Dehaene, S., & Changeux, J.-P. (1993). Development of Elementary Numerical Abilities: A Neuronal Model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407. https://doi.org/10.1162/jocn.1993.5.4.390

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How Does the Brain Solve Visual Object Recognition? *Neuron*, *73*(3), 415–434. https://doi.org/10.1016/j.neuron.2012.01.010

Drew, T., & Vogel, E. K. (2008). Neural Measures of Individual Differences in Selecting and Tracking Multiple Moving Objects. *Journal of Neuroscience*, *28*(16), 4183–4191. https://doi.org/10.1523/JNEUROSCI.0556-08.2008

Egeth, H. E., Leonard, C. J., & Palomares, M. (2008). The role of attention in subitizing: Is the magical number 1? *Visual Cognition*, *16*(4), 463–473. https://doi.org/10.1080/13506280801937939

Ester, E. F., Drew, T., Klee, D., Vogel, E. K., & Awh, E. (2012). Neural Measures Reveal a Fixed Item Limit in Subitizing. *Journal of Neuroscience*, *32*(21), 7169–7177. https://doi.org/10.1523/JNEUROSCI.1218-12.2012

Ester, E. F., Klee, D., & Awh, E. (2014). Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1022–1033. https://doi.org/10.1037/a0035377

Ester, E. F., Zilber, E., & Serences, J. T. (2015). Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of Vision*, *15*(1), 4–4. https://doi.org/10.1167/15.1.4

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Foster, D. H., Gravano, S., & Tomoszek, A. (1989). Acuity for fine-grain motion and for two-dot spacing as a function of retinal eccentricity: Differences in specialization of the central and peripheral retina. *Vision Research*, *29*(8), 1017–1031.

Francis, G., Manassi, M., & Herzog, M. H. (2017). Neural dynamics of grouping and segmentation explain properties of visual crowding. *Psychological Review*, *124*(4), 483–504. https://doi.org/10.1037/rev0000070

Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113*(1), 1–13. https://doi.org/10.1016/j.cognition.2009.07.002

Gallistel, C. R., & Gelman, R. (1991). Subitizing: The preverbal counting process. *Memories, Thoughts, and Emotions: Essays in Honor of George Mandler*, 65–81.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*(1–2), 43–74. https://doi.org/10.1016/0010-0277(92)90050-R

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences*, *4*(2), 59–65. https://doi.org/10.1016/S1364-6613(99)01424-2

Goldfarb, L., & Levy, S. (2013). Counting within the Subitizing Range: The Effect of Number of Distractors on the Perception of Subset Items. *PLoS ONE*, *8*(9), e74152. https://doi.org/10.1371/journal.pone.0074152

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences*, *106*(31), 13130–13135.

Greenwood, J. A., Sayim, B., & Cavanagh, P. (2014). Crowding is reduced by onset transients in the target object (but not in the flankers). *Journal of Vision*, *14*(6), 2–2. https://doi.org/10.1167/14.6.2

Grill-Spector, K., & Sayres, R. (2008). Object Recognition: Insights from Advances in fMRI Methods. *Current Directions in Psychological Science*, *17*(2), 73–79.

Harrison, W. J., & Bex, P. J. (2015). A Unifying Model of Orientation Crowding in Peripheral Vision. *Current Biology*, *25*(24), 3213–3219. https://doi.org/10.1016/j.cub.2015.10.052

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334–337. https://doi.org/10.1038/383334a0

Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: a fresh look at object recognition. *Current Opinion in Behavioral Sciences*, *1*, 86–93. https://doi.org/10.1016/j.cobeha.2014.10.006

Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *Journal of Vision*, *15*(6), 5. https://doi.org/10.1167/15.6.5

Intriligator, J., & Cavanagh, P. (2001). The Spatial Resolution of Visual Attention. *Cognitive Psychology*, *43*(3), 171–216. https://doi.org/10.1006/cogp.2001.0755

Jevons, W. S. (1897). The power of numerical discrimination. *Nature*, *3*, 281–282. https://doi.org/doi:10.1038/003281a0

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The Discrimination of Visual Number. *The American Journal of Psychology*, *62*(4), 498. https://doi.org/10.2307/1418556

Kennedy, G. J., & Whitaker, D. (2010). The chromatic selectivity of visual crowding. *Journal of Vision*, *10*(6), 15–15. https://doi.org/10.1167/10.6.15

Keshvari, S., & Rosenholtz, R. (2016). Pooling of continuous features provides a unifying account of crowding. *Journal of Vision*, *16*(3), 39. https://doi.org/10.1167/16.3.39

Kleiner, M., Brainard, D., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1.

Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A Shared, Flexible Neural Map Architecture Reflects Capacity Limits in Both Visual Short-Term Memory and Enumeration. *Journal of Neuroscience*, *34*(30), 9857–9866. https://doi.org/10.1523/JNEUROSCI.2758-13.2014

Koffka, K. (1935). Principles of Gestalt theory. *New York: Hartcourt*.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*(2), 255–279.

Korte, W. (1923). Über die Gestaltauffassung im indirekten Sehen. *Zeitschrift Für Psychologie*, *93*, 17–82.

Krajcsi, A., Szabó, E., & Mórocz, I. Á. (2013). Subitizing Is Sensitive to the Arrangement of Objects. *Experimental Psychology*, *60*(4), 227–234. https://doi.org/10.1027/1618-3169/a000191

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*(5), 635–654. https://doi.org/10.1016/j.visres.2007.12.009

Levi, D. M., & Carney, T. (2009). Crowding in Peripheral Vision: Why Bigger Is Better. *Current Biology*, *19*(23), 1988–1993. https://doi.org/10.1016/j.cub.2009.09.056

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, *2*(2), 3–3.

Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision*, *7*(2), 4–4.

Logan, G. D., & Zbrodoff, N. J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review*, *10*(3), 676–682.

Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision*, *7*(2), 24. https://doi.org/10.1167/7.2.24

Manassi, M., & Whitney, D. (2018). Multi-level Crowding and the Paradox of Object Recognition in Clutter. *Current Biology*, *28*(3), R127–R133.

Mandler, G., & Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1.

Mareschal, I., Morgan, M. J., & Solomon, J. A. (2010). Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision*, *10*(8), 13–13.

Marr, D. (1982). *Vision: A computational investigation into*. WH Freeman.

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, *5*(1), 6. https://doi.org/10.1167/5.1.6

Mazza, V., & Caramazza, A. (2015). Multiple object individuation and subitizing in enumeration: a view from electrophysiology. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00162

Mazza, V., Pagano, S., & Caramazza, A. (2013). Multiple Object Individuation and Exact Enumeration. *Journal of Cognitive Neuroscience*, *25*(5), 697–705. https://doi.org/10.1162/jocn_a_00349

Mazza, V., Turatto, M., & Caramazza, A. (2009). Attention selection, distractor suppression and N2pc. *Cortex*, *45*(7), 879–890. https://doi.org/10.1016/j.cortex.2008.10.009

Nandy, A. S., & Tjan, B. S. (2007). The nature of letter crowding as revealed by first- and second-order classification images. *Journal of Vision*, *7*(2), 5. https://doi.org/10.1167/7.2.5

Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nature Reviews Neuroscience*, *6*(3), 177–190. https://doi.org/10.1038/nrn1626

Palomares, M., Smith, P. R., Pitts, C. H., & Carter, B. M. (2011). The effect of viewing eccentricity on enumeration. *PLoS One*, *6*(6), e20779.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*(7), 739–744.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12), 12. https://doi.org/10.1167/4.12.12

Pelli, D. G., & Tillman, K. A. (2007). Parts, wholes, and context in reading: A triple dissociation. *Plos One*, *2*(8), e680.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 1129–1135. https://doi.org/10.1038/nn.2187

Petrov, Y., & Meleshkevich, O. (2011). Asymmetries and idiosyncratic hot spots in crowding. *Vision Research*, *51*(10), 1117–1123. https://doi.org/10.1016/j.visres.2011.03.001

Piazza, M., & Eger, E. (2016). Neural foundations and functional specificity of number representations. *Neuropsychologia*, *83*, 257–273. https://doi.org/10.1016/j.neuropsychologia.2015.09.025

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, *121*(1), 147–153. https://doi.org/10.1016/j.cognition.2011.05.007

Piazza, M., Mechelli, A., Butterworth, B., & Price, C. J. (2002). Are Subitizing and Counting Implemented as Separate or Functionally Overlapping Processes? *NeuroImage*, *15*(2), 435–446. https://doi.org/10.1006/nimg.2001.0980

Poder, E. (2006). Crowding, feature integration, and two kinds of “attention”. *Journal of Vision*, *6*(2), 7–7.

Roelfsema, P. R., & Houtkamp, R. (2011). Incremental grouping of image elements in vision. *Attention, Perception, & Psychophysics*, *73*(8), 2542–2572. https://doi.org/10.3758/s13414-011-0200-0

Rosen, S., Chakravarthi, R., & Pelli, D. G. (2014). The Bouma law of crowding, revised: Critical spacing is equal across parts, not objects. *Journal of Vision*, *14*(6), 10–10. https://doi.org/10.1167/14.6.10

Sayim, B., & Wagemans, J. (2017). Appearance changes and error characteristics in crowding revealed by drawings. *Journal of Vision*, *17*(11), 8–8.

Scolari, M., Kohnen, A., Barton, B., & Awh, E. (2007). Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays? *Journal of Vision*, *7*(2), 7–7.

Shin, K., Chung, S. T. L., & Tjan, B. S. (2017). Crowding, visual awareness, and their respective neural loci. *Journal of Vision*, *17*(5), 18–18. https://doi.org/10.1167/17.5.18

Strasburger, H., & Malania, M. (2013). Source confusion is a major cause of crowding. *Journal of Vision*, *13*(1), 24–24. https://doi.org/10.1167/13.1.24

Toet, A., & Levi, D. M. (1992). The Two-Dimensional Shape of Spatial. Retrieved from http://www.academia.edu/download/41763762/toet_levi1992.pdf

Trick, L. M. (1987). Subitizing and canonical pattern matching. In *Annual meeting of the Canadian Psychological Association, Vancouver, Canada*.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*(1), 80.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of Face Perception. *Annual Review of Neuroscience*, *31*(1), 411–437. https://doi.org/10.1146/annurev.neuro.30.051606.094238

Tyler, C. W., & Likova, L. T. (2007). Crowding: a neuroanalytic approach. *Journal of Vision*, *7*(2), 16–16.

Ullman, S. (1996). *High-level vision: Object recognition and visual cognition* (Vol. 2). MIT press Cambridge, MA.

Ullman, S. (2007). Object recognition and segmentation by a fragment-based hierarchy. *Trends in Cognitive Sciences*, *11*(2), 58–64. https://doi.org/10.1016/j.tics.2006.11.009

van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2010). A Neurophysiologically Plausible Population Code Model for Feature Integration Explains Visual Crowding. *PLoS Computational Biology*, *6*(1), e1000646. https://doi.org/10.1371/journal.pcbi.1000646

Vanrullen, R. (2009). Binding hardwired versus on-demand feature conjunctions. *Visual Cognition*, *17*(1–2), 103–119. https://doi.org/10.1080/13506280802196451

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating Attentional Load Affects Numerosity Estimation: Evidence against a Pre-Attentive Subitizing Mechanism. *PLoS ONE*, *3*(9), e3269. https://doi.org/10.1371/journal.pone.0003269

Vickery, T. J., Shim, W. M., Chakravarthi, R., Jiang, Y. V., & Luedeman, R. (2009). Supercrowding: Weakly masking a target expands the range of crowding. *Journal of Vision*, *9*(2), 12–12. https://doi.org/10.1167/9.2.12

Wender, K. F., & Rothkegel, R. (2000). Subitizing and its subprocesses. *Psychological Research*, *64*(2), 81–92.

Whitney, D. (2005). Motion distorts perceived position without awareness of motion. *Current Biology : CB*, *15*(9), R324–R326. https://doi.org/10.1016/j.cub.2005.04.043

Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*(4), 167–174. https://doi.org/10.1016/j.tics.2009.01.008

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*(3), 295–340.