Vision Research 122 (2016) 34–42

Contents lists available at ScienceDirect

Vision Research

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

Numerosity underestimation in sets with illusory contours

Atanas Kirjakovski, Eriko Matsumoto *

Graduate School of Intercultural Studies, Kobe University, 1-2-1 Tsurukabuto, Nada-ku, Kobe 657-8501, Japan

A R T I C L E I N F O

Article history:
Received 20 October 2015
Received in revised form 16 March 2016
Accepted 18 March 2016
Available online 6 April 2016

Keywords:
Numerosity estimation
Segmentation
Object perception
Configural processing
Illusory contours

A B S T R A C T

People underestimate the numerosity of collections in which a few dots are connected in pairs by task-irrelevant lines. Such configural processing suggests that visual numerosity depends on the perceived scene segments, rather than on the perceived total area occupied by a collection. However, a methodology that uses irrelevant line connections may also introduce unnecessary distraction and variety, or obscure the perception of task-relevant items, given the saliency of the lines. To avoid such potentially confounding variables, we conducted four experiments where the line-connected dots were replaced with collinear inducers of Kanizsa-type illusory contours. Our participants had to compare two simultaneously presented collections and choose the more numerous one. Displays comprised c-shaped inducers and disks (Experiment 1), c-shaped inducers only (Experiments 2 and 4), or closed inducers (Experiment 3). One display always showed a 12- (Experiments 1–3) or 48-item reference pattern (Experiment 4); the other was a test pattern with numerosity varying between 9 and 15 (Experiments 1–3) or 36–60 items (Experiment 4). By manipulating the number of illusory contours in the test patterns, the level of connectedness increased or decreased respectively. The fitted psychometric functions revealed an underestimation that increased with the number of illusory contours in Experiments 1 and 2, but was absent in Experiments 3 and 4, where illusory contours were more difficult to perceive or larger numerosities were used. Results corroborate claims that visual numerosity estimation depends on segmented inputs, but only within moderate numerical ranges.

1. Introduction

The debate on how numerosities are extracted from visual scenes endures in the literature, oscillating between those who suggest that numerosity estimation operates over unsegmented visual inputs (Allik & Tuulmets, 1991; Durgin, 2008), and those who think that scene segmentation is obligatory (Castelli, Glaser, & Butterworth, 2006; Dehaene & Changeux, 1993). At the heart of the problem lies the dilemma of whether objects in collections retain the properties of the single objects upon which numerosity estimation operates, or whole collection representations are formed, averaged and analyzed in a global manner. In favor of the unsegmented inputs view, there is evidence that the physical size of a set can be estimated by a single mean value calculated from all its members, without their compulsory identification (Ariely, 2001). Furthermore, numerosity is confounded with other attributes like texture density and area (Durgin, 1995; Hollingsworth, Simmons, Coates, & Cross, 1991; Vos, van Oeffelen, Tibosch, & Allik, 1988), frame/area ratio of array displays (Bevan & Turner, 1964) and dot cluster, which shows an inverse relation with the estimates of its numerosity; that is, sets with clustered items tend to be underestimated (Ginsburg & Goldstein, 1987; Frith & Frit, 1972). A suitable theoretical framework for such observations is given by Allik and Tuulmets (1991) who proposed the *occupancy model*, describing numerosity as an emergent property from a summed area occupied by all items in a set (for an extension of the model that combines texture density and area, see Durgin, 1995). In other words, numerosity estimation relies on a single statistic, called *an occupancy index*, calculated over the whole collection. Later, a different model of numerosity estimation was proposed by Dehaene and Changeux (1993), in which object individuation and segmentation, rather than occupied areas, serve as the basis for the estimation. The information about objects is gathered by an "input retina" and forwarded to an intermediate module that normalizes and encodes all objects regardless of their size, as segmented activation units with Gaussian distributions. In support of this view, studies on crowding show that although the object recognition is severely impaired in cluttered scenes, detection of whole objects, or even an object's individual features, remains unaffected (Whitney & Levi, 2011). In other words, the signal of individual objects does not disappear

---

* Corresponding author.
E-mail addresses: kirjakovski@gmail.com (A. Kirjakovski), ermatsu@kobe-u.ac.jp (E. Matsumoto).

http://dx.doi.org/10.1016/j.visres.2016.03.005
0042-6989/

even in cluttered visual scenes, giving the opportunity for the numerosity extractors to operate over discrete units. It seems that the Dehaene and Changeux (1993) model does not account for the effects of clustering, covered area, and object size reported previously; but Franconeri, Bemis, and Alvarez (2009) hypothesize that such factors could interact with numerosity extraction during the later stages of responding (i.e., after the number estimates have already been acquired).

One way to test whether visual numerosity estimation depends on segmented inputs is to change the level of *connectedness* between the separate items while keeping the spatial factors constant across the display. That is, by connecting items in a way that is irrelevant to the number task (e.g., lines), it is possible to increase or decrease the perceived segmentation of items in a configuration, since line-connected objects tend to group together and are perceived as a single unit rather than as two independent objects (Palmer & Rock, 1994). If numerosity estimation is performed over unsegmented inputs, connecting individual items into pairs should not influence the total estimation of a set. In contrast, if segmented units are the basis for the estimation, the perceived numerosity should vary inversely with the level of connectedness inside the collection (i.e., the more connected pairs a set contains, the less numerous it should appear). Franconeri et al. (2009) and He, Zhang, Zhou, and Chen (2009) confirmed the latter hypothesis by demonstrating that people systematically underestimated the total number of items when these were presented as line-connected pairs. The underestimation was absent when participants in both studies compared disconnected items, with lines attached only to a single item or freely hanging among the rest of the items in the collection. However, Franconeri et al. (2009) and He et al. (2009) reached their conclusions from experiments in which connectedness was achieved solely by means of physical connections. In such experimental designs, lines inherently become distractors, since participants are asked to ignore them and only estimate the numerosity of other elements. Yet, previously it has been shown that simply instructing participants to ignore irrelevant stimuli does not abolish the effects of distractors. For example, a distractor's proximity to a target is a significant factor when people make responses to the specified target; that is, the closer the distractor is, the bigger the interference with the target (Eriksen & Eriksen, 1974). The interference is particularly strong and unavoidable when distractors and targets both belong to the same object (Kramer & Jacobson, 1991). Additionally, lines potentially affect the salience of connected items tagged for enumeration, and the reported underestimation could be simply due to a reduced visibility of the individual items in the pair rather than—as the authors explained—a perceptual grouping into unified objects.

In the current study, we tried to avoid any confounding effects of the connecting lines by adapting the method described in He et al. (2009) using Kanizsa illusory contours (ICs) instead. By definition, ICs are perceived boundaries and edges in the absence of physical correlates of luminance, color and textural changes in the stimulus (Kanizsa, 1976). They are extremely versatile and robust visual illusions that are perceived by human adults and infants (Bertenthal, Campos, & Haith, 1980), as well as animals, including birds and insects (Nieder, 2002). ICs initiate neural responses early in the visual process, as early as those of the V1 and V2 areas in monkeys and humans (Ffytche & Zeki, 1996; Grosof, Shapley, & Hawken, 1993; von der Heydt, Peterhans, & Baumgartner, 1984). Critically, the pattern of activity in neurons from the V2 area, when their receptive fields fall within the empty gap between IC inducers, resembles that previously found in anesthetized cats when light-bars are shone in a particular orientation (Hubel & Wiesel, 1962; Peterhans & von der Heydt, 1989). Such properties of the ICs make them an ideal candidate for replacing physical lines but still maintaining different levels of connectedness between the items in a collection.

We carried out four experiments in which we asked people to identify the numerically larger set from two briefly-shown panels on a computer screen. One of the panels always contained a fixed numerosity reference set, while the other was a test set with the number of items varying between 9 and 15 in Experiments 1, 2, and 3, and between 36 and 60 in Experiment 4. The stimulus sets in Experiment 1 consisted of both disks and c-shaped inducers (the latter subset giving rise to the ICs), whereas sets in Experiments 2 and 4 contained only inducers. Experiment 3 was similar to Experiment 2, except that the inducers were closed with a thin line. Experiment 4 was a modified version of Experiment 2 that investigated numerosities four times larger. Test sets across experiments contained 0, 2 or 4 ICs in Experiments 1–3, and 0, 8, and 16 ICs in Experiment 4, each formed by a subset of items. Similar to Franconeri et al. (2009) and He et al. (2009), we reasoned that if our visual system extracts numerosity from segmented and unitized inputs, there should be an underestimation in test sets with more ICs, due to the mandatory binding of the inducing elements into unified objects. Conversely, underestimation should not be evident in Experiment 3 because closing the gaps of inducing items disrupts the completion of ICs (Peterhans & von der Heydt, 1991).

2. Experiment 1: inducers and disks

2.1. Description

The first experiment adapted the study of He et al. (2009) by replacing the physical links with Kanizsa-type ICs. Participants were not informed about ICs in the test patterns, and confusion about the concept of numerosity as it relates to connecting lines was clarified by instructing them to use all presented stimuli.

2.2. Materials and methods

2.2.1. Participants

Six people—two males (ages 30 and 31 years, including one of the authors) and four females (ages 24, 24, 25, and 27 years)—participated in the first experiment. Excluding the author, all other participants were naïve regarding the purpose of the experiment. In this and subsequent experiments, participants had normal or corrected-to-normal vision, were rewarded with a USB flash drive (monetary value of 800 JPY) for their participation, and were internationally, racially, and culturally diverse. In all experiments, informed consent was obtained from all participants, and the study was conducted in accordance with the Declaration of Helsinki.

2.2.2. Stimuli, design, and procedure

2.2.2.1. Stimuli. The experiment was conducted on a Mac mini computer (OS X 10.8.5) with FlexScan L568 EIZO 17″ monitor (338 mm × 270 mm active display size and 1280 × 1024 pixels resolution). At a viewing distance of 80 cm where participants were seated, 1 pixel subtended a visual angle of 0.0187°. The stimulus patterns were generated in a custom Python program and displayed using PsychoPy software (Peirce, 2007). The total stimulus collection contained 168 reference and 168 test patterns, each uniquely and randomly generated offline. The reference patterns were constructed from four black disks (diameter = 20 pixels; RGB = 0, 0, 0) and eight misaligned c-shaped inducers (diameter = 20 pixels; notch width = 4 pixels; and notch length = 10 pixels, measured from the center of the inducer outwards; RGB = 0, 0, 0), with a constant number of 12 items per reference pattern (see Fig. 1). The reference patterns did not contain ICs, and their absence was confirmed through a visual inspection of