When The Weaker Conquer: A Contrast-Based Illusion of Visual Numerosity and Its Dependence on Segregation

by Quan Lei

B.S. in Psychology, Central China Normal University
M.S. in Psychology, Peking University

A dissertation submitted to

The Faculty of
the College of Science of
Northeastern University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

June 9, 2015

Dissertation directed by

Adam Reeves
Professor of Psychology
Acknowledgements

At this moment of closing my dissertation work, I’m grateful to all these people who have made this achievement possible.

First of all, I'd like to sincerely thank my dear advisor, Prof. Adam Reeves. Although I hesitated when making the decision about where to pursue my PhD degree four years ago, I have never regretted my decision ever since. I feel so lucky to have been here at Northeastern and worked together with this great man. I can remember that at the beginning when I showed little interest in the project I was involved in, he patiently explained to me the significance of the work and pointed me to new potential directions over and over again. He would sit in my office for hours and talk to me with ideas for experiments and inspirations of any random sort. It is through these numerous conversations that I truly believe that Adam has treated me like a colleague and friend with much more respect than I should have deserved as a student. Our collaboration has been very enjoyable but not without stories. When I was desperately searching for a dissertation topic to work on, I seemed to be motivationally lost for some periods of time. Adam noticed this and was worried about my progress, but instead of confronting me in person, he wrote me a harsh email warning me to shape up. On the other hand, he carried out all the daily conversations with me just like nothing had happened, still patiently guiding me towards potential dissertation projects. In this process, he has been extremely understanding and open-minded. If not for that, I couldn’t have ended up working on something that doesn’t really fit into his research interest. We have then enjoyed working together on a totally new topic, which has culminated in this dissertation. The experience
working with Adam, with all his passion for science and rigorousness in mind, is a great fortune to me, which is certain to continue enlightening me for the years to come.

I also want to sincerely thank Prof. Rhea Eskew and Prof. Stephen Harkins for being on my dissertation committee. Rhea has been very critical but at the same time extremely supportive. This dissertation has benefitted a lot from his detailed comments and all the discussions that have happened between us. It is always a joy to talk to him, who has made the process of developing a dissertation much less stressful. I’m also appreciative of Rhea as a super delightful person and I have enjoyed sharing the same section of the building with him. I’d like to extend a special thanks to Steve for embarking on an unusual journey with me by sitting on my committee. Without much personal connection, he willingly accepted to get involved in my research that is hardly relevant to his own area. He has been so devoted and has provided me with invaluable advice on improving my dissertation. He would go into the details and spot inconsistencies I couldn’t have noticed by myself. I really appreciate your help, Steve. With regard to the dissertation, I also want to thank Prof. Peter Bex for being a reader and getting back to me with very helpful comments.

I wish to thank the Psychology Department at Northeastern for enabling me the opportunity to pursue my PhD degree in such a vibrant and lovely atmosphere. I have happily spent my years here with the support and care from the staff and the faculty. I especially want to thank Prof. Judith Hall for all her effort in making my life easier as a graduate student. Judy is like a good friend who is always available to help me with any
questions I have. She is so attentive and our interaction has always been filled with joy. I’m also lucky to have made friends with many fellow graduate students in the department. In particular I’d like to thank Timothy Shepard, Jeffrey Nador, and Jiehui Qian for sharing research interest and for accompanying me during this endeavor. You have been a large part of my life here at Northeastern. I especially want to thank Jiehui for great friendship and tremendous intellectual and emotional support when I need it even after her graduation from Northeastern.

Last but certainly not least, I want to thank my family for their enduring love. I’m grateful to my parents, who themselves are not well educated but are so much dedicated to my education for so many years. I’m also grateful to my brother and my sister-in-law for their endless support and my little nieces for readily welcoming me home. All your love has been the impetus for my success.
Abstract of Dissertation

Humans and many non-human species alike are endowed with a capability to perceive the approximate numerosity of a large set. In the visual domain, various factors have been found to affect perceived numerosity. These include continuous extent variables (i.e., length, area, density, etc.), spatial structure of the stimulus (e.g. clustering) and visual contexts. In the current study, the effect of luminance contrast on approximate numerosity perception was investigated.

A numerosity illusion was found in which low-contrast elements appear to be more numerous than high-contrast ones when they were intermixed in the same display. Using a numerosity discrimination task, the illusion, with an average magnitude of ~10%, was observed with either positive (white and light gray disks on a dark field) or negative contrast (black and dark gray disks on a light field). The illusion was eliminated when gray and white stimuli were segregated by visual field, motion (whites moved left when grays moved right, and vice-versa), orientation (using white and gray vertical and horizontal bars) or stereoscopic depth. Therefore, the illusion might be due to a failure in segregation between the two sets of elements.

When matching the numerosity of gray or white disks in an intermixed patch to the numerosity of an homogenous patch (gray or white disks unmixed), the numerosity of white disks in the intermixed patch was underestimated relative to the homogenous patch, but the numerosity of gray disks was veridically matched, evidence of an asymmetrical interaction between the two disk sets in an intermixed patch, such that the stronger (white)
elements appeared less numerous. The same results were obtained when the irrelevant (not-to-be-matched) set of disks in the intermixed patch was replaced by visual noise of the same contrast, refuting explanations of the illusion in terms of misclassification (whites being seen as grays more often than the reverse) or occlusion (whites hiding potential grays).

Lastly, it was also shown that the numerosity illusion was not specific to the numerosity discrimination task but also occurred for absolute numerosity estimation. The reported numerosity of gray disks in an intermixed patch was consistently higher than that of white ones, regardless of how subjects allocated their attention between the two disk sets, which points to a pre-attentive locus of the effect.

To conclude, when unsegregated, an asymmetrical sensory/perceptual interaction occurs automatically between low-contrast and high-contrast elements, resulting in an illusory reduction in the numerosity of high-contrast ones. To account for the illusion, a thresholding process that operates at an early stage of visual processing to separate the two sets of disks for enumeration is hypothesized.
Table of Contents

Acknowledgements ii

Abstract of Dissertation v

Table of Contents vii

List of Figures ix

Chapter 1: General Introduction 1
  1.1 A Preverbal Number System 1
  1.2 Numerosity Apprehension: One System or Two? 6
  1.3 The Direct vs. Indirect Controversy 11
  1.4 Factors in Numerosity Judgment 15
  1.5 The Number Sense: Neural Mechanisms 28
  1.6 Computational Models of Numerosity 33

Chapter 2: Study Overview 41

Chapter 3: General Methods 44
  3.1 Subjects 44
  3.2 Stimuli 44
  3.3 Psychometric Procedures 45

Chapter 4: Quantifying the Numerosity Illusion 49
  4.1 Experiment 1: Gray vs. White Reference 49
  4.2 Experiment 2: Negative Contrast 52

Chapter 5: The Role of Segregation 55
  5.1 Experiment 3: Segregation by Visual Field 55
  5.2 Experiment 4: Segregation by Stereoscopic Depth 59
  5.3 Experiment 5: Segregation by Motion 63
  5.4 Experiment 6: Segregation by Orientation 66

Chapter 6: Revealing the Interaction: Matching Experiments 75
  6.1 Experiment 7: Superset Matching 76
  6.2 Experiment 8: Gray/White in Intermixture vs. G/W in Isolation 79
  6.3 Experiment 9: Matching in Noise 82

Chapter 7: Absolute Numerosity Estimation 86
  7.1 Experiment 10a: Post-cue Random Design – Intermixed 88
  7.2 Experiment 10b: Pre-cue Random Design – Intermixed 91
  7.3 Experiment 10c: Block Design – Intermixed 93
  7.4 Experiment 10d: Unmixed Condition 94
  7.5 Discussion 96
Chapter 8: General Discussion
  8.1 Mechanism of the Illusion and Numerosity Computation 99
  8.2 Saliency/Attention, Perceptual Grouping and Numerosity 108
  8.3 Contextual Effect in Numerosity Perception 111
  8.4 The Role of Contrast in Visual Processing 113

References 118
List of Figures

Figure 1. Numerosity adaptation. 12
Figure 2. Numerosity illusions. 21
Figure 3. Connectedness and numerosity. 23
Figure 4. Numerosity selectivity in neurons. 29
Figure 5. A neuronal model of numerosity. 35
Figure 6. Area models of numerosity. 38
Figure 7. Example stimulus and psychometric function. 45
Figure 8. Example stimulus in a matching task. 47
Figure 9. Result of Experiment 1: white vs. gray reference. 51
Figure 10. Example stimulus and result of Exp. 2: negative contrast. 53
Figure 11. Example stimulus and result of Exp. 3: segregation by visual field. 57
Figure 12. Example stimulus and result of Exp. 4: depth segregation. 61
Figure 13. Example stimulus and result of Exp. 5: motion segregation. 64
Figure 14. Example stimuli and results of Exp. 6: segregation by orientation. 71
Figure 15. Result of Exp. 7: superset matching. 78
Figure 16. Result of Exp. 8: intermixture vs. isolation. 80
Figure 17. Example stimulus and result of Exp. 9: matching in noise. 83
Figure 18. Flow of a numerosity estimation trial in Experiment 10. 89
Figure 19. Results of Exp. 10: numerosity estimation. 91
Figure 20. The thresholding model. 104
Chapter 1. General Introduction

Life is impossible without numbers, everyone would agree. Number is indeed an integral part of our daily life. When we go grocery shopping, everything is quantified by numbers. We may purchase one box of chocolate, two bottles of milk, three pounds of apples and so on. We regularly pay bills that are expressed in numbers. We also use numbers to describe time and space, like in ‘it took me 45 minutes to get to work’ and ‘I walked 4 miles today’. When we go to school, we study mathematics, which is the subject that deals with numbers. Mathematics provides the basis for pursuing science and technology, the advancement of which has characterized our modern society. Needless to say, numerical skills have been well developed in humans, especially within advanced human cultures. Despite the highly sophisticated form modern mathematics may take, it cannot be built upon nothing. Where does all this come from? According to some, underlying our ability to create and understand mathematics is a number sense or mathematical instinct that is shared by human adults, human infants and even other animal species (Dehaene, 2011).

1.1 A Preverbal Number System

While advanced types of numerical capacities such as exact calculation are believed to be uniquely human and supported by the language system, some rudimentary forms of numerical competence have been demonstrated in non-human animals and also in human infants, suggesting the existence of an evolutionarily old and developmentally early system for processing numerical information that operates independently of the language system. It is not too difficult to imagine the survival values a numerical competence may
afford to animals living in a natural environment. In a fight or flight situation, individuals may base their decisions on the relative number of their friends and foes. In foraging activity it is also important to decide which tree bears more fruit and is worth more time spent picking. Indeed, it has been demonstrated in playback experiments that female lions upon hearing the roars of extragroup intruders adjusted their approaching behavior according to the relative size of the intruding group and their own group (McComb, Packer, & Pusey, 1994). Free-ranging rhesus monkeys have also been found to make spontaneous numerical judgment when choosing apple slices from several containers in experiments simulating a foraging task (Hauser, Carey, & Hauser, 2000). Comparable numerical competence has also been demonstrated in animal species that are phylogenetically even further removed from humans, including in birds (Scarf, Hayne, & Colombo, 2011), in fish (Agrillo, Dadda, Serena, & Bisazza, 2009) and in insects (Dacke & Srinivasan, 2008).

Remarkable numerical competence has also been documented in human infants well before the onset of the verbal system. Infants as young as five to six month old can discriminate between visual arrays comprising different numbers of objects (Starkey & Cooper, 1980; Xu & Spelke, 2000). Newborn infants can determine the numerical equivalence between visuospatial arrays of objects and temporal sequences of auditory events (Izard, Sann, Spelke, & Streri, 2009), and 7-month-old infants spontaneously associate visual displays of adult faces with the matching number of voices they hear speaking (Jordan & Brannon, 2006). More strikingly, infants as young as 5 months can even perform elementary arithmetic, both addition and subtraction (Wynn, 1992). Infants
in this study watched the experimenter acting out either correct or false arithmetical operations by adding or removing toys on a stage with the aid of a hiding screen. They were consistently surprised and looked longer when a violation of the expected result was acted out, as in ‘1+1=1’ and ‘2-1=2’, compared to the correct operations of ‘1+1=2’ and ‘2-1=1’. These findings strongly argue for a preverbal number system that works its way towards higher levels of numerical skills yet to come later in life.

Further evidence supporting the existence of a preverbal number system came from studies of numerical competence in some indigenous human cultures. Members of the Amazonian tribe Piraha speak a language lacking in words for numbers beyond two. While they were found to be severely limited in their ability to enumerate exact quantities, when they were asked to construct sets of nuts and batteries to match an example set in number, their performance exhibited the signature of an analog estimation process that followed Weber law, i.e. a constant coefficient of variation (Gordon, 2004), which was also observed in other human adults (Whalen, Gallistel, & Gelman, 1999) and in other animal species (Platt & Johnson, 1971). Another indigenous culture extensively studied is Munduruku, also residing in the Amazonian region and using a language with words only for numbers one through five. Although the members of Munduruku had difficulty performing exact arithmetic, their performance on large number comparison and approximate calculation with dot arrays closely matched that of a French group used as the control, again obeying the Weber law and showing the numerical distance effect (to be described in the next paragraph)(Pica, Lemer, Izard, & Dehaene, 2004). These studies have revealed a dissociation in numerical capacities: exact enumeration (usually
involving verbal counting) and calculation which both depend intimately on the language system, and an analog estimation process that operates independently of language and is shared across cultures and species. It is the latter that is believed to constitute our mathematical instinct and found to correlate with mathematical achievement in humans (Halberda, Mazzocco, & Feigenson, 2008).

Whereas we use number symbols on a daily basis to represent the quantity of objects, the core representations upon which the symbolic number system builds are believed to be nonsymbolic in nature. Moyer and Landauer (1967) provided the earliest evidence that numbers are represented in an analogue format similar to continuous physical magnitude such as length and loudness. They asked subjects to judge pairs of numerals from 1-9 and decide which one of each pair was larger. Reaction time (RT) in this task was found to be an inverse function of the numerical distance between the two numerals to be compared (1 vs. 9 faster than 1 vs. 5), hence the numerical distance effect. When the distance was constant, RT was faster when the numerals were smaller (1 vs. 2 faster than 2 vs. 3), hence a magnitude effect. Taken together, time required for making a numerical decision depends on the ratio between the two numerals in each pair, suggesting that symbolic numbers are represented as analogue magnitude that obeys Weber’s law. This proposition has received support from other studies. Whalen et al. (1999) engaged subjects in a speeded key press task in which they had to press the key for a specified number of times, while suppressing them from verbally counting the number of presses. It was found that both mean response and standard deviation of response were proportional to the required number of key presses, such that their ratio, i.e. the coefficient of variation, remained
constant across the range of numbers (7-25) tested. A constant coefficient of variation was also observed when subjects were asked to estimate the number of flash events. This result replicated that already found in rats engaging in the same key-press task (Platt & Johnson, 1971), suggesting that humans share with other animal species an evolutionarily primitive system that represents number as analogue magnitude with scalar variability. It has been demonstrated that this system operated almost indistinguishably in monkeys and in humans when subjects from both species performed a number ordering task, showing the same ratio dependence for both accuracy and reaction time (Cantlon & Brannon, 2006). Ratio dependence of numerical judgment is also evident in human infants. While 6-month-old infants successfully discriminated 4 vs. 8 or 8 vs. 16, they failed to discriminate 4 vs. 6 or 8 vs. 12 (Wood & Spelke, 2005; Xu & Spelke, 2000).

While acknowledging the existence at least in humans of a symbolic number system that stores numerical knowledge and supports verbal counting, this dissertation will focus on the process of apprehending numbers through the preverbal number system. While ‘number’ is a generic concept by which we refer to the quantity of something and is usually associated with the symbolic system, the term ‘numerosity’ will be preferred when we refer to the discrete quantity of a set of objects, such as a dot array or a sequence of sound pulses, although sometimes these two terms will still be used interchangeably in this dissertation. Despite the fact that studies on numerosity have looked into multiple sensory modalities (Barth, Kanwisher, & Spelke, 2003), the major focus of this dissertation will be on visual numerosity.
1.2 Numerosity Apprehension: One System or Two?

Despite the vast evidence that we possess a preverbal system that represents number in an analogue fashion, it may not be a unitary one. The situation is complicated by the observation that numerical judgment about extremely small numbers (1-4) is so precise and effortless that they appear to be special compared to large numbers. It is now widely held that there are two systems underlying our general numerical competence: one for representing large and approximate numerosities, the approximate number system, and a second one for representing small and exact numerosities, or the exact number system (Feigenson, Dehaene, & Spelke, 2004). More than one century ago, Jevons (1871) empirically measured the capacity for accurately enumerating beans thrown into a box. Using himself as the subject, he concluded that he could enumerate at most four items without errors in an instant. This is perhaps the earliest empirical study on what is later called ‘subitizing’, a rapid and mostly errorless enumeration process with a capacity limit of about four to six items (Kaufman, Lord, Reese, & Volkmann, 1949). Within the subitizing range, enumeration performance is nearly perfect and reaction time is essentially flat with a slope of about 40-100ms per item; beyond this range, both error rate and reaction time rise abruptly, when counting is involved, at a cost of 250-350ms with each additional item (Atkinson, Campbell, & Francis, 1976). This discontinuity in curve slope has led many to believe that there exists a specialized system for enumerating small numerosities up to about four items. When enumerating more than four items without verbal counting, a numerosity estimation process is assumed to take place. Then what causes the discontinuity between small and large numerosities?
One explanation posits that a small number of items in most cases form canonical patterns and recognition of these patterns facilitates enumeration (Mandler & Shebo, 1982). According to this hypothesis, for example, one can easily associate a line pattern formed by two dots with the number 2 and associate a triangular pattern formed by three dots with the number 3. These are naturally occurring canonical patterns. It was demonstrated that, with some training, the exploitation of canonical patterns could lead to perfect enumeration beyond the typical subitizing range. Another explanation for the superior subitizing performance resorts to a limited-capacity preattentive mechanism in visual processing, based on the so-called FINST (for FINgers of INSTantiation) theory (Trick & Pylyshyn, 1994). According to this theory, only a limited number of items can be indexed or individuated in parallel preattentively by the visual system and subsequently accessed for enumeration, hence the subitizing capacity. Beyond this limit, serial spatial attention is required for counting the items. Consistent with this hypothesis, it was shown that manipulating spatial attention had a greater effect on counting latency than on subitizing latency.

Some other researchers, however, suggest that subitizing is nothing more than just a special case of the approximate number system and therefore advocate the idea of one unitary system. Balakrishnan and Ashby (1992) analyzed the distribution of reaction time (RT) data from a speeded enumeration task using linear arrays of items and found no discontinuity in RT data as numerosity increased. Instead, response time required for enumeration showed a continuous increase with each additional item, even within the typical subitizing range. Cordes, Gelman, Gallistel, and Whalen (2001) demonstrated that
in a key press task the coefficient of variation, which is the ratio of standard deviation to mean response, was constant across the whole range of numerosities tested, both small and large, suggesting that a common mechanism underlies the representation of numerosities both within the putative subitizing range and beyond. van Oeffelen and Vos (1982) suggested that the same Weber fraction applies to the discrimination between any two numerosities, including those within the subitizing range. They measured the Weber fraction in a numerosity discrimination task and estimated it to be .162, which roughly corresponds to the difference between 6 and 7. According to their probability model, the nearly perfect performance within the subitizing range is due to the fact that any numerosity within this range is readily discriminable from any other numerosities because their difference is guaranteed to be above the discriminable difference dictated by the fixed Weber fraction. Ross (2003) held the same view, although the Weber fraction measured by him was somewhat higher at .25, which corresponds to the difference between 4 and 5.

Revkin, Piazza, Izard, Cohen, and Dehaene (2008) argued that, if the same estimation mechanism that operates on large numerosities also operates with subitizing but with higher precision due to Weber law, one would expect similarly perfect performance for discriminating between large numerosities that have the same ratio as that between small numerosities within the subitizing range. They trained subjects with decade numerosities of 10 through 80 and then tested them in a numerosity naming task. Against the prediction from Weber law, subitizing-like performance was not observed for this series of numerosities. Compared to the series of 1 through 8, naming responses to the 10-80
series were also much more variable, as reflected in elevated variation coefficients, suggesting different mechanisms for different ranges of numerosities.

Evidence supporting separate systems for representing small and large numerosities also came from infant studies. Using a preferential looking paradigm, Xu (2003) demonstrated that 6-month-old infants could successfully discriminate 4 from 8 elements in a dot array but failed to discriminate 2 from 4, despite the same ratio for the two pairs. It is assumed that infants have a capacity limit of 3 items for representing small numbers of objects and it is difficult, if not impossible, for infants to make numerical comparison across this boundary when two separate systems are implicated (Cordes & Brannon, 2009a). The same failure was replicated when infants had to discriminate the number of jumps of a puppet (Wood & Spelke, 2005). Feigenson, Carey, and Hauser (2002) reported a similar failure: when 10 to 12-month-old infants were allowed to choose from two containers with three or fewer crackers in both of them, the infants consistently chose the one with more crackers; however, when one container had more than three crackers in it and the other less, the infants chose randomly from the two. This pattern was also found in non-human animals (Hauser et al., 2000), suggesting that there exists an evolutionary basis for two separate systems for representing small and large numerosities.

Another line of research, rather than directly bears on the one vs. two systems controversy, has challenged the preattentive nature of subitizing as proposed by Trick and Pylyshyn (1994). It is generally found that diverting attention away from a numerical task degrades performance even for small numerosities within the subitizing range. Olivers
and Watson (2008) combined a letter identification task and a dot enumeration task in an attentional blink paradigm. Except for the smallest numerosity which was one, enumeration of all other numerosities (2-5) was subject to the deteriorating influence of attentional blink (see similar result in Egeth, Leonard, & Palomares, 2008). More interestingly, when the enumeration task preceded the identification task, the effect of attentional blink on identification accuracy depended on the numerosity of the dot array so that performance in the identification task dropped more as the numerosity increased, suggesting a graded attentional demand for enumeration within the subitizing range. Subitizing is not only susceptible to the manipulation of temporal attention as in attentional blink but also to the diversion of spatial attention. Vetter, Butterworth, and Bahrami (2008) compared subjects’ enumeration performance under three conditions: single task, dual task with low load, in which the primary task was a feature detection task, and dual task with high load, in which the primary task was a conjunction detection task. An orderly decrease in enumeration accuracy with increasing attention load was observed for all numerosities 1 through 8 and the decrease was even more pronounced for small numerosities within the traditional subitizing range than beyond. In several other studies that employed a similar dual task paradigm, it was all found that enumeration accuracy for small numerosities suffered from the depletion of attentional resources, in sharp contradiction to the preattentive theorization of subitizing (Burr, Turi, & Anobile, 2010; Railo, Koivisto, Revonsuo, & Hannula, 2008).

These results have led researchers to reconsider the controversy of one vs. two systems for representing small and large numerosities. One emerging view holds that the
approximate number system operates on both small and large numerosities, which is complemented by an attention-dependent mechanism with high precision underlying the representation of small numerosities within the subitizing range. This explains why enumeration performance with small numerosities becomes more similar to that of large numerosities when attentional resource is limited. Consistent with this view, Burr, Anobile, and Turi (2011) demonstrated that numerosity adaption affected the perceived numerosities in the subitizing range under high attentional load but not otherwise.

1.3 The Direct vs. Indirect Controversy

Burr and Ross (2008) advocated numerosity as a unique visual attribute that deserves a mechanism of its own, on a par with those well-studied visual dimensions of color, motion and orientation. They based this conclusion on the observation that numerosity is reliably adaptable, with other aspects of the stimulus well controlled. In their paradigm, they adapted the subjects to either a large number or a small number of dots and then measured the effect of adaptation on the apparent numerosity of a test stimulus (Fig. 1). It was discovered that adapting to a large numerosity reduced and adapting to a small numerosity increased the apparent numerosity of the test stimulus, in the best case by a factor of 2. More importantly, when they mismatched the adaptor and the test stimulus in element size (or pixel density), contrast and orientation, the adaptation effect was hardly affected. They concluded that it is numerosity that is adapted, rather than other covarying factors. There is evidence that adaptation takes place across a wide range of numerosity, including the subitizing range if the attentional load is high (Burr et al., 2011). Because adaptation has been regarded by psychophysicists as a diagnostic tool for the primacy of
a visual dimension, they proposed that the visual system can sense numerosity directly without the mediation of surrogate measures. As a primary visual attribute, numerosity obeys Weber’s law (Ross, 2003).

Figure 1. Numerosity adaptation. Upper panel: adaptor stimuli with a dense (high-numerosity) dot patch on the left and a sparse (low-numerosity) patch on the right; lower panel: test stimuli with two dot patches of equal numerosity. Fixate the red spot in the upper panel for 30s and then shift fixation to the red spot in the lower panel. The left patch should now appear to be much less numerous than the right patch. (Burr & Ross, 2008)
However, this idea has been challenged by some researchers who hold that numerosity can only be inferred from other surrogate measures, one of which they advocated is texture density. Durgin (2008) argued that the numerosity adaptation effect reported by Burr and Ross (2008) is actually the consequence of texture density adaptation. In an earlier study, it was discovered that adapting to a dense dot texture pattern reduced both the perceived density and the perceived numerosity of a subsequent dot pattern (Durgin, 1995). However, while the density aftereffect was constant across numerosity levels of the test pattern, the numerosity aftereffect was range-dependent: it was commensurate with the density aftereffect only at high numerosities and fell short of it at low numerosities. To account for this discrepancy, it was found that the perceived clustering of the test dot pattern increased after adaptation, which combined with the density change determined the perceived numerosity. By this logic, numerosity is not sensed directly but inferred from other sources of information, texture density being particularly relevant at high numerosities. Dakin, Tibber, Greenwood, Kingdom, and Morgan (2011) furthered this idea by proposing a model that computes density by taking the ratio of the response of two spatial filters, one tuned to high frequency and the other to low frequency content of an input image. Numerosity is then gauged by scaling the response ratio with the patch size of a dot pattern. This model explicitly instantiates the idea that numerosity is derived from texture density and is used by Dakin et al. (2011) to explain their empirical observation that a larger patch of dots appears to be both denser and more numerous than a smaller patch that physically matches the larger one in either density or numerosity, with the bias being more pronounced for density than for numerosity. Consistent with the prediction of this model, another study found that numerosity and density discrimination
were similarly affected by mismatching the two dot patches to be compared in patch size, element size, contrast and by manipulation of attention, suggesting a common underlying metric (Tibber, Greenwood, & Dakin, 2012).

Ross and Burr (2010) argued that if numerosity is mediated by density, since additional noise will be introduced to derive numerosity from density, the precision of numerosity discrimination should be lower than that of density discrimination, and numerosity judgment should suffer when density is not a reliable predictor of numerosity. They compared conditions where the area, density or numerosity of a dot pattern was kept constant while the other two factors covaried, these conditions either blocked or intermingled. No evidence was found that the Weber fraction for numerosity was larger than that for density, suggesting that numerosity discrimination is at least as precise as density discrimination. The Weber fraction for numerosity under the constant density condition was decreased compared to the constant area condition when all conditions were intermingled, which is opposite to the prediction made if numerosity judgment is mediated by density. Although this study has provided some evidence against the hypothesis that numerosity is computed through density, it does not necessarily imply that numerosity is instead sensed directly. It is still possible that the visual system draws on various sources of information to infer numerosity and texture density may represent one important source. Actually, numerosity judgment was found to benefit from density information when the density of a dot pattern was relatively high (Anobile, Cicchini, & Burr, 2014). In this study, subjects performed either numerosity or density discrimination between two patches of dots with equal or unequal area. It was found that, under the
equal area condition, the Weber fraction for numerosity discrimination was constant at low numerosities (densities) but decreased steadily once the density of the display surpassed a certain value (~0.3 dots/deg$^2$), at the same rate as the decrease in the Weber fraction for density. To explain the coincidental drop in threshold at high densities for both numerosity and density discrimination, the authors suggested that, when the two patches were equal in area, density could be strategically used as a reliable cue to numerosity. Under the unequal area condition where this strategy was discouraged, however, the Weber fraction for numerosity was constant throughout and didn’t follow that of density. Based on these results, the authors advocated separate mechanisms for numerosity and density. Can numerosity be sensed directly? Given the conflicting evidence, the jury is still out.

1.4 Factors in Numerosity Judgment

1.4.1 Continuous Variable vs. Discrete Numerosity

It is intuitive to think that numerosity is a highly abstract property of a visual set, the extraction of which appears impossible without the aid of verbal counting. In natural environment, objects for enumeration come in all varieties. With regard to numerosity, it doesn’t even matter whether it is one apple or one elephant when we refer to the number ‘one’. One apple can be red or green and can be big or small. One hundred apples can either be scattered in an apple tree or collected in piles for sale in a supermarket. Clearly, to perceive numerosity with at least some precision, one has to disregard various visual properties of a visual set that usually coexist/covary with but nevertheless irrelevant to numerosity. Despite ample evidence that humans and other species alike possess an
incredible capacity for apprehending numerosity, one particular concern in the literature is to what extent this capacity relies on the exploitation of continuous extent variables (i.e., length, area, density, etc.) in a visual set. This concern is a legitimate one, since in laboratory studies it is usually difficult to cleanly disentangle numerosity from continuous properties of the artificial stimuli used. Indeed, there has been a huge controversy, particularly in the domain of infant cognition, as to whether numerosity can be represented at all if continuous extent variables are controlled.

Clearfield and Mix (1999) pitted total contour length of elements against numerosity in a preferential looking paradigm. They habituated 6 to 8-month old infants to visual displays of either two or three squares and then tested them with two types of display: one with the familiar number of squares but a novel total contour length and the other with the familiar total contour length but a novel number. They found that infants dishabituated to the visual display with a novel total contour length but not to the display with a novel number of squares, leading them to conclude that numerosity is not spontaneously represented by infants and if they discriminate numbers at all, they do so on the basis of continuous extent properties like contour length rather than of discrete number itself. Infants were also found to attend to total area of elements rather than the number of them when contour length was controlled (Clearfield & Mix, 2001), suggesting a general sensitivity of infants to spatial extent variables over numerosity. This finding has also been extended to three-dimensional objects. It was found that infants kept track of the total front surface area of the sets of 3-D objects but not the number (Feigenson, Carey, & Spelke, 2002). Similarly, in a cracker choice task, infants
appeared to base their choices on the total volume rather than the number when the crackers were of different sizes (Feigenson, Carey, & Hauser, 2002). Xu, Spelke, and Goddard (2005) also reported that when the total filled area of the sets of objects were controlled for infants failed at discriminating between small numerosities (1 vs. 2). All these results suggest that infants are quite able to compute the extent properties of a visual set and usually make decisions on the basis of these properties; when these properties correlate with the number of items in the set, it gives the ‘false’ impression that numerosity is spontaneously extracted and used by the infants. Is this a fair claim? Perhaps not!

Cordes and Brannon (2009b), in an attempt to replicate the study by Clearfield and Mix (1999), found that infants dishabituated to both change in total contour length and change in numerosity for small sets of objects. When large sets were involved, however, infants only noticed the change in numerosity but not change in continuous extent variable (either contour length or area) when both underwent a 2-fold change. Brannon, Abbott, and Lutz (2004) also reported that infants failed to attend to a change in continuous extent variable (cumulative area) when the numerosity was controlled. Furthermore, it was shown that a 4-fold change in cumulative surface area was required for infants to notice the change, much higher a ratio than that required for noticing a change in numerosity (Cordes & Brannon, 2008). These results have challenged the view that numerosity is not as salient a dimension of any visual set as certain continuous extent variables and suggest instead that it is actually much easier for infants to keep track of numerosity than those continuous variables.
How can these conflicting pieces of evidence reconciled? There appears to be a dissociation between small and large sets: infants are more sensitive to continuous variables than to numerosity or equally sensitive to both when small sets of objects are involved, but numerosity apparently wins over continuous variables when large sets are involved. It appears that computing the continuous properties of a visual set is only viable for a small number of objects but the extraction of numerosity is relatively robust across both small and large sets of objects. Whether one relies on numerosity or continuous variables may also depend on their relative saliency in a specific task. For example, infants attend to cumulative area in a 1 vs. 2 discrimination task when the objects involved were similar to each other (Feigenson, Carey, and Spelke, 2002) but they attend to numerosity with heterogeneous objects (Feigenson, 2005). Presumably, it makes more sense for infants to summate a continuous extent variable across several homogenous objects than across heterogeneous ones for which tracking each individual seems more relevant, hence the differential observations.

Assuming numerosity is indeed one visual dimension which we can apprehend with some degree of precision, then to what extent is numerosity judgment influenced by continuous perceptual variables, such as item size, aggregate surface and stimulus area (or convex hull)? Miller and Baker (1968) reported that small items were overestimated and large items underestimated by their subjects in a numerosity estimation task. Such an inverse relation between size and numerosity has also been established in several other studies (Gebuis & Reynvoet, 2012a, 2012b; Ginsburg & Nicholls, 1988; Catherine Sophian,
However, Hurewitz, Gelman, and Schnitzer (2006) reported that item size interferes with numerosity judgment in a way that subjects were biased to judge the display consisting of larger items as more numerous, that is, numerosity judgment was more accurate when item size was congruent with the numerosity of a stimulus. What about aggregate surface? Aggregate surface refers to the actual area taken up by a set of elements, which is the sum of the area of individual elements. Hurewitz et al. (2006) again reported a congruency effect between this variable and numerosity: subjects showed a tendency to judge the stimulus with a larger aggregate surface area to be more numerous. By contrast, Gebuis and Reynvoet (2012a, 2012b) found an inverse relation between this variable and numerosity. While aggregate surface simply sums up the area of individual elements in a set, stimulus area refers to a contiguous portion of the background over which the elements of a set are distributed. It is usually operationalized as convex hull, the smallest convex boundary that can be delineated to enclose all elements of a set. Krueger (1972) reported that subject gave a higher estimation for the numerosity of an array of dots spread over a larger area than the same number of dots over a smaller area. This effect of stimulus area or convex hull on numerosity judgment has been replicated a number of times by other researchers (Dakin et al., 2011; Gebuis & Reynvoet, 2012a, 2012b; Sophian & Chu, 2008; Tokita & Ishiguchi, 2010b; but see Birnbaum, Kobernick, & Veit, 1974; Birnbaum & Veit, 1973). While the evidence we have so far are still inconclusive about exactly how each of these continuous variables are taken into account in numerosity judgment, it is quite clear that such a judgment is not made independently of these non-numerical factors coexisting with numerosity in a stimulus. The tight correlation of many of these variables (e.g. area
and density) poses a great methodological challenge for numerosity studies: it is difficult to manipulate a variable without affecting another (Gebuis & Reynvoet, 2011). To the extent that numerosity judgment is achieved by weighing a number of variables which may have opposing effects, this difficulty and associated methodological differences across studies may explain some of the inconsistencies in the findings.

Many studies investigating the effect of continuous variables on numerosity judgment have adopted a congruency paradigm to probe into this question (Gebuis & Reynvoet, 2012a; Hurewitz et al., 2006; Sophian, 2007; Sophian & Chu, 2008). In this paradigm, numerosity is pitted against one or more continuous variables so as to create two trial conditions: in the congruent condition, the stimulus (usually an array of dots) with a higher (lower) numerosity value also has a larger (smaller) value in the continuous variable in question; in the incongruent condition, the stimulus higher (lower) in numerosity is instead smaller (larger) in that continuous variable. If a particular continuous variable plays a role in numerosity judgment, a congruency effect should be expected. However, even if a continuous variable exerts some consistent bias on numerosity judgment, it is not necessarily inconsistent with the existence of an approximate numerosity system that extracts numerosity independent of other magnitude information, because any bias observed may reflect a competition at the cognitively driven response stage rather than an early-on intrusion into numerosity representation at the perceptual level. This point of view is supported by the fact that in most studies even when numerosity is incongruent with the continuous variable, subjects were still able to perform numerosity discrimination well above the chance level. Continuous variables
like those mentioned above can actually be exploited by the visual system to refine numerosity judgment, since in natural situations many of these variables are highly correlated with numerosity.

![Numerosity illusions. (A) Solitaire illusion. The filled (black) circles appear to be more numerous than the empty circles, although they are of equal numbers (Frith & Frit, 1972). (B) Regular-random illusion. The regular dot pattern on the left appears to be more numerous than the random pattern on the right, although there are equal number of dots in both patterns (Ginsburg, 1976).](image)

### 1.4.2 Spatial Arrangement/Structure

Frith and Frit (1972) described the so-called solitaire illusion (Fig. 2A). In the illustration, filled circles appear to be more numerous than empty ones, although there are equal numbers of them. Of course whether a circle is filled or empty is irrelevant here and is just for illustrative purpose. According to Frith and Frit (1972), what matters is the Gestalt each set of circles form. The filled circles because of their contiguity form a good Gestalt that leads to a higher perceived numerosity, compared to the spatially separated empty circles. Consistent with the Gestalt principle, number estimation was also found to benefit from pattern symmetry, one Gestalt determinant of good visual organization.
(Howe & Jung, 1987). These effects are believed to occur automatically at the perceptual level.

Ginsburg (1976) described a numerosity illusion that involves comparing a regular pattern vs. a random pattern (see Messenger, 1903, for an earlier mention). The regular pattern in Fig. 2B was judged to be more numerous than the random pattern although they contain the same number of dots, an illusion of ~5.5% in magnitude. It was further found that subjects overestimated the numerosity of a regular pattern and underestimated that of a random pattern relative to their physical number, which still held when the covered area (convex hull) of the patterns were matched (Ginsburg, 1978, 1980). Apparently, the distribution of items within a pattern matters in numerosity judgment. In an attempt to quantify the different distributions of these patterns, Ginsburg and Goldstein (1987) proposed the cluster continuum, onto which each pattern can be positioned. The degree of clustering of a pattern was indexed by cluster ratio. To derive this measure, the whole pattern is divided into a number of equally-sized cells. After the count of items in each cell is known, we can get the mean number of items per cell, M and also the variance, V, the ratio between which gives the cluster ratio, V/M. Conceivably, a regular pattern has a cluster ratio of zero and as the items within a pattern clump together, the ratio increases. A true random pattern by definition has a ratio of one, following the Poisson distribution. Another index to clustering was also proposed: the mean distance to nearest neighbor, $d_{nn}$ (Ginsburg, 1991). This distance decreases as the items within a pattern become more clustered, with a regular pattern highest in this value. With either measure, there appears to be a quantifiable relationship between the
perceived numerosity of a pattern and its cluster level: the perceived numerosity decreases as the cluster level increases (Ginsburg, 1991; Ginsburg & Goldstein, 1987; Sophian & Chu, 2008). Therefore, the regular vs. random illusion is just a special case following from this general rule.

A

B

Figure 3. Connectedness and numerosity. (A) Dots for enumeration are linked into polygons (Koesling, Carbone, Pomplun, Sichelschmidt, & Ritter, 2004). (B) Squares for enumeration are connected with lines to form cube-like objects (Franconeri, Bemis, & Alvarez, 2009).

Which items within a pattern cluster together is determined to a large extent by spatial proximity. One way to override this type of grouping is to explicitly connect items using connectors like lines, the so-called uniform connectedness principle (Palmer & Rock, 1994). How are connected items enumerated? Koesling, Carbone, Pomplun, Sichelschmidt, and Ritter (2004) connected dots into polygons (Fig. 3A) and this manipulation led to severe underestimation of the numerosity of dots in the display compared to a display with the same number of dots but no connecting lines. In a similar vein, Franconeri et al. (2009) connected pairs of squares into cube-like units (Fig. 3B).
and asked subjects to discriminate the numerosity of squares in such a display from that in a display where those cube-like units were cut in halves so that each individual square still had irrelevant lines attached but were not connected to other squares. It turned out that the ‘connected’ display was perceived to be much less numerous than the ‘unconnected’ display. This underestimation in numerosity of connected items persisted even when a much less prominent connector, i.e., a single thin line, was used to connect pairs of circles (see also He, Zhang, Zhou, & Chen, 2009, for similar methods and results). These studies have important implications as to what are the basic units for enumeration. The hypothesis here is that salient objects as defined by uniform connectedness in these cases can’t be ignored. It seems that these task-irrelevant objects automatically compete with task-relevant objects (or sub-objects) for numerosity representation and the resulting perception seems to be a compromise between the two. Consistent with this view, visually nesting objects so that smaller objects were contained within bigger ones was also found to reduce the perceived numerosity of the whole set (Chesney & Gelman, 2012).

Depth order was also found to bias numerosity judgment. Schutz (2012) discovered that, in a random-dot kinematogram (RDK) display, dots with a certain motion direction that were perceived to lie on a back surface were judged to be more numerous than dots perceived to be on a front surface. When binocular disparity was added to override motion direction-induced depth ordering, compared to a single-surface reference, there was a consistent bias for subjects to overestimate the numerosity of dots on the back surface and underestimate the numerosity of dots on the front surface. Tsirlin, Allison,
and Wilcox (2012), though using a very different stimulus, i.e. densely packed small dots that clearly formed a texture, also discovered that the numerosity (or density) of elements stereoscopically assigned to a back depth plane was overestimated relative to elements assigned to a front plane. They explained this phenomenon with a model in which higher-level processes of figure-ground assignment feedback to disparity-tuned neurons favoring those with the same disparity as the ground which is usually the back depth plane. By this account, perceived numerosity is intimately related to neural signal strength whatever this may be.

1.4.3 Contextual Effect

Enumeration tasks in real world situations are usually carried out within a certain context. We may want to estimate the number of passengers on a subway train or the number of spectators at a sporting event. More often than not the objects of interest for enumeration don’t occur alone but coexist with currently uninteresting objects or distractors. For example, one may want to estimate the number of fans of a sport team who are interspersed among fans of the opponent team. Despite this fact, studies on numerosity perception have generally used much artificial stimuli consisting of isolated sets of objects, whether they were designed to have subjects estimate the numerosity of a single set or discriminate numerosity between two or more sets. Contextual influences seem to be the law rather than exception in visual processing (Albright & Stoner, 2002). Our perception of motion, color and much more is clearly modulated by various contextual factors. There is some evidence showing that numerosity judgment is also subject to contextual influences.
In an early study, subjects were asked to estimate the number of beans in a glass jar (Bevan, Maier, & Helson, 1963). They were either instructed to view the jar as part of the figure stimulus that went together with the beans or view it as background separate from the beans. It was found that subjects gave higher estimation for beans in a large jar than for beans in a small jar, hence an assimilation effect, but only when the jar was viewed as part of the figure. When it was viewed as the background, there was no effect of jar size. In a similar study, subjects estimated the numerosity of a dot array bounded by either a large or a small frame and the figure-ground relationship between the dots and the frame was again manipulated by giving different instructions to different groups (Bevan & Turner, 1964). An assimilation effect was found in the figure condition so that numerosity estimation was higher for dots bounded by a large frame than for dots bounded by a small one, replicating the result of Bevan et al. (1963). However, a contrast effect was also observed in the ground condition: estimation was lower for dots bounded by a larger frame than for those bounded by a smaller one. These studies have provided some early evidence that numerosity estimation is vulnerable to the influence of contextual factors.

More recently, Halberda, Sires, and Feigenson (2006) asked a somewhat different question: what is our capacity for enumerating sets of objects that are spatially overlapping? To answer this question, their stimulus consisted of one to six sets of differently colored dots that were randomly intermixed. On each trial, subjects had to estimate the number of the dot set that shared a particular color (the color subset) or the
number of all the dots in the patch (the superset) regardless of color. In the probe-before condition, subjects knew in advance which color subset (or superset) to estimate on each trial, while in the probe-after condition, subjects were not informed of the target set until after the stimulus had disappeared. They found that the cost associated with the probe-after condition, as reflected in the magnitude of estimation error, was only evident for color subset estimation when there were three or more subsets involved in the dot array. However, superset estimation didn’t differ between probe-before and probe-after conditions for any number of sets included in the array. The authors concluded that when enumerating spatially intermixed objects there is a capacity limit: we can simultaneously enumerate three sets of objects: two (color) subsets and the superset. This capacity limit could reflect a limit on numerosity processing, but could also be due to a similar limit on visual working memory for colors in that particular stimulus setup (Poltoratski & Xu, 2013).

Cordes, Goldstein, and Heller (2014) also used stimuli that consisted of differently colored dots that belonged to two subsets and were spatially intermixed. In an attempt to investigate how the presence of one subset (distractor subset) affects enumeration of the other subset (target subset), they varied the numerosity of the distractor subset as: no distractor; fewer distractors than targets; same number of distractors as targets; and more distractors than targets. The estimated numerosity of the target subset decreased as the size of the distractor subset increased in both probe-before and probe-after conditions. This result was interpreted by the authors as a simultaneous contrast effect which is presumed to operate pre-attentively. This is an interesting finding and clearly shows that,
like what is already well known in many other visual processes, contextual effects also occur in the numerosity domain.

1.5 The Number Sense: Neural Mechanisms

What neural substrates underlie the numerical competence of humans and other species as reviewed above? Or more specifically, how is numerosity represented in the brain? Sawamura, Shima, and Tanji (2002) demonstrated that a large proportion of neurons in superior parietal lobule were selective to the number of movement executions when the experimental monkeys were engaged in a sensorimotor task in which two types of movements alternated after five repetitions of each. It appeared that individual neurons were able to ‘count’ the number of movement executions in a behaviorally relevant manner. However, representation of sensorimotor numerosity by these neurons is not abstract in the sense that it was found to be specific to the type of movement, whether it’s a ‘push’ or ‘turn’.

Nieder and colleagues (Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2004) discovered that, with other confounding stimulus parameters controlled, many neurons in the prefrontal cortex (PFC) and posterior parietal cortex (PPC) responded to a visual display comprising a small number of dots in a way that demonstrated numerosity tuning. Monkeys were trained to perform a delayed match-to-quantity task (Fig. 4A) in which they had to either release or hold a lever depending on whether a test stimulus matched the sample stimulus in numerosity, while single-cell activity is recorded in a population of neurons at PFC (Nieder et al., 2002) and PPC (Nieder & Miller, 2004). It was found
that 30% of neurons in lateral PFC and 18% of neurons in the fundus of the intraparietal sulcus (IPS) showed selectivity for numerosity (Fig. 4B), each having its own preferred numerosity. Each of these neurons showed peak response to its preferred numerosity and the spiking rate systematically dropped off as the numerical distance between a test
numerosity and the preferred numerosity increased, which mimics the well-known numerical distance effect found with human subjects (Moyer & Landauer, 1967). Furthermore, neurons in PPC had a shorter latency of numerosity selectivity than neurons in PFC, suggesting that numerosity information arises first in PPC and flows to PFC therefrom. When directly comparing behavioral performance and neuronal activity (Fig. 4C and 4D), filter functions of both were symmetrical on a logarithmic scale, consistent with the Weber-Fechner law, suggesting that numerosity is represented in the form of analog magnitude (Nieder & Miller, 2003). This type of numerosity tuning was also found to exist in PFC for large numerosities up to 30 (Nieder & Merten, 2007).

Converging evidence supporting numerosity tuning came from fMRI studies with human subjects. Piazza, Izard, Pinel, Le Bihan, and Dehaene (2004) employed a fMRI adaptation paradigm, in which subjects passively viewed a stream of dot patterns with a fixed numerosity (16 or 32) as the habituation stimuli and after initial adaptation a range of numerosities deviant from the habituation numerosity were sparsely inserted among the habituation sets. Using this paradigm, fMRI signal recovery was observed for the deviant numerosities in both left and right intraparietal sulci (IPS), with the recovery amount positively dependent on the numerical distance between the deviant and the habituation numerosity. Activation in both IPSs was nicely fit by an inverted Gaussian function of the ratio between the deviant and habituation numerosity on a logarithmic scale, consistent with data both from separate behavioral tasks in the same study and from single-unit recording in primate animals (Nieder & Miller, 2003). Is numerosity tuning in the parietal lobe specific to numerosity conveyed by dot patterns? At least one
study suggests no. Piazza and colleagues (Piazza, Pinel, Le Bihan, & Dehaene, 2007) in another study used the same fMRI adaptation paradigm to investigate whether numerosities conveyed by dot patterns and by numerical symbols (Arabic digits) share a common magnitude representation. During fMRI scanning, the habituation stimuli could either be Arabic digits or dot patterns and the sparse deviant stimuli could also be in either format. It was found that fMRI signal recovered in a numerical distance-dependent fashion irrespective of the specific way of conveying numerosity, so that recovery was also evident for deviant dot patterns after adapting to Arabic digits and vice versa, suggesting a common underlying magnitude representation. Therefore, both single-unit neurophysiological studies in primate animals and fMRI studies in humans have revealed a labeled-line magnitude code for numerosity in parietal regions of the brain.

However, a labeled-line code may not be the only coding scheme for numerosity. Roitman, Brannon, and Platt (2007) trained monkeys to perform an implicit numerosity discrimination task while recording from a population of neurons in the lateral intraparietal (LIP) region of the brain. The LIP neurons studied have contralateral receptive fields to which a dot pattern with either a standard or a deviant numerosity was presented. However, the dot patterns were task-irrelevant as they only predicted the amount of reward the monkeys would get while engaging in a saccadic task in which they had to make a saccade to a target on the opposite side to the numerosity stimulus upon the extinguishing of the fixation point. A large number of LIP neurons were found to respond to the dot pattern within their receptive fields in a graded fashion, following a monotonic function of the number of dots in the stimulus. The firing rate of some neurons
kept rising as the numerosity increased while that of others showed the opposite trend. This response profile of LIP neurons is in sharp contrast to the tuning properties of PFC and IPS neurons studied by Nieder and colleagues (Nieder et al., 2002; Nieder & Merten, 2007; Nieder & Miller, 2004). These results are not necessarily incompatible, though. As the authors suggest, these LIP neurons with a graded coding scheme may serve as magnitude accumulation units which provide input to the numerosity-tuning neurons found in other regions of the parietal lobe, as proposed by some computational models (Dehaene & Changeux, 1993; Meck & Church, 1983). While this may be true for LIP neurons with a monotonically increasing response, it remains unclear what functions those neurons showing a monotonically decreasing response to numerosity may serve. Consistent with the single neuron findings, a brain region in the posterior superior parietal cortex was found to increase in activation with increasing numerosity in an fMRI study with human subjects, suggesting summation coding rather than tuning (Santens, Roggeman, Fias, & Verguts, 2010).

So far, abundant evidence points to the posterior parietal cortex (PPC) as a key site in the brain for numerosity representation. This notion has also received support from lesion studies (Cipolotti, Butterworth, & Denes, 1991; Lemer, Dehaene, Spelke, & Cohen, 2003) and a TMS study (Cappelletti, Barth, Fregni, Spelke, & Pascual-Leone, 2007), showing that both permanent and temporary dysfunction of the parietal lobe (affecting the PPC region) disrupts numerical processing, particularly when approximation of numerosity was involved. Although these studies have generally implicated PPC in numerosity representation, little is known about the spatial structure of these representations. With
respect to this aspect, Eger et al. (2009) demonstrated that, using multivariate analysis of fMRI data, individual numerosities could be decoded from brain activation patterns in the intraparietal sulcus (IPS). More recently, using population receptive field analysis, Harvey, Klein, Petridou, and Dumoulin (2013) revealed that in the posterior superior parietal lobule there exists a representation map for small numerosities which is organized topographically, so that from the medial to the lateral part of the region of interest (ROI) there is a progressive increase in preferred numerosity. Consistent with other studies (Piazza et al., 2004; Whalen et al., 1999), tuning width increased with preferred numerosity in an approximately linear fashion. It was also found that more cortical surfaces were devoted to representing lower than higher numerosities, resonating with the similar observation that in the prefrontal cortex of monkeys more neurons preferred lower than higher numerosities (Nieder & Merten, 2007). To summarize, numerosity tuning, combined with topographic mapping, suggests that numerosity is represented in the parietal cortex in much the same way as more basic visual dimensions such as orientation, spatial frequency and motion are represented in the visual cortex (De Valois, Albrecht, & Thorell, 1982; De Valois, Yund, & Hepler, 1982; Zimmermann et al., 2011).

1.6 Computational Models of Numerosity

A number of models have been developed to help understand how numerosity information is extracted by both humans and other animals. In its simplest form, Meck and Church (1983) proposed the mode control model. According to this model, number counting and timing share the same mechanism relying on an accumulation process. The
mechanism can run in either the ‘event’ mode to accumulate number of events (or objects) from the pacemaker or in the ‘run’ and ‘stop’ mode to accumulate elapsed time. The core idea of this model is that numerosity, like time interval, is represented by analogue magnitude, which is corroborated by both psychophysical and neuroimaging findings (Moyer & Landauer, 1967; Piazza et al., 2004). Several other models, though much more elaborated over this one, have adopted a similar accumulation mechanism as an intermediate step between sensory input and numerosity output.

The model by Dehaene and Changeux (1993) is made up of four stages of processing (Fig. 5). Each input object is first coded as a local Gaussian distribution of various sizes by a sheet of neuron clusters on a topographically organized ‘retina’. The second stage consists of neuron clusters that function as difference-of-Gaussian (DoG) filters on the sensory input. These DoG filters shift in receptive field location along one dimension and vary in widths along another, each responding preferentially to an object falling within its receptive field and having a size that matches its filter width. This layer essentially conducts size normalization so that a certain number of input objects, regardless of variation in sizes, activate approximately the same number of clusters on a location map. Activity from this stage is then pooled by a layer of summation units with differential thresholds, each responding when the pooled activity, which correlates highly with the numerosity of input objects, exceeds a certain limit. As can be seen, activation pattern in this layer has already become sensitive to numerosity. To further derive numerosity selectivity or tuning, however, these summation units project topographically onto a layer of numerosity detectors in a way that each detector receives excitatory input from its
corresponding summation unit but lateral inhibition from others. The consequence is that each numerosity detector develops tuning properties, responding preferentially to a cardinal numerosity and declining in activation with increasing distance from this preferred numerosity. It should be noted that all the connections within this model are preset manually, with the implicit assumption that there exists a hard-wired neural network devoted to numerical processing, in accordance with the idea that humans and other species share an innate primitive number system (Dehaene, 2011).

![Neuronal Model of Numerosity](image)

**Figure 5.** A neuronal model of numerosity (Dehaene & Changeux, 1993).

To the contrary, other network models of numerosity take a learning approach. Verguts and Fias (2004) devised a backpropagation model with three fields: an input location field, a hidden field and an output number field. Each unit in the hidden field receives a linear combination of the inputs from units in the location field. After transformation with
thresholding, activities in the hidden field are then fed into the number field, again linearly combined and transformed. All connections within the network were initially set with random weights. They trained the model so that the correct unit in the number field will be activated in response to an input numerosity. The interesting finding is that, after training, summation coding emerges in the hidden layer so that each hidden unit pools activities from the input layer and responds monotonically to numerosity, similar to the property of the summation units in the model by Dehaene and Changeux (1993). Therefore, an analog summation process appears to be a natural intermediate step bridging the sensory input and numerosity output. Using this summation field as the input layer, which connects to the number field, they further trained an unsupervised learning model without specifying any of the connections between the two fields. After Hebbian-like learning, each unit in the number field ends up spontaneously tuning to a specific numerosity. This result suggests that numerosity is a learnable dimension of environment stimuli, the extraction of which can be implemented by an initially uncommitted network after some learning. In a similar vein, Stoianov and Zorzi (2012) showed that a number ‘sense’ can emerge within a deep learning network. Their hierarchical generative model consists of a visible layer encoding the input image and two hidden layers that learn to efficiently code the sensory input or in another sense to reconstruct it through top-down connections. After unsupervised learning, some units in the second hidden layer spontaneously represent local numerosity by combining input from two types of neurons: classical center-surround neurons and another type of neurons that encode the cumulative area of stimulus and provide a normalization signal. Both types of neurons emerge in the first hidden layer. Population activity of these numerosity detectors showed monotonic
coding of numerosity and was found to support numerosity comparison with a precision comparable to that of human subjects. The successful extraction of numerosity information by this model corroborates the view that numerosity is a learnable feature of sensory images and a pre-designated network for numerical processing may not be necessary. It is worth noting that all these models, despite their specific forms, point to a summation stage in the build up of numerosity representation. The neural underpinnings of the summation mechanism have been revealed by both single neuron recording in monkeys and neuroimaging in humans (Roitman et al., 2007; Santens et al., 2010).

Other models, however, on the assumption that numerosity is not itself directly perceptible, tried to circumvent the whole concept of a numerosity representation and to explain numerosity judgment in terms of surrogate measures. Vos, van Oeffelen, Tibosch, and Allik (1988) proposed the filled area of a pattern (in their case a dot array) as the index to its numerosity. The first step in their model is to determine the contour of a pattern using the CODE (from COntour DEtector) algorithm. They posited that where the subjective contour lies depends on the interaction between each dot and its neighbors. A spread function is applied to each dot with a dispersion equaling half the distance of this dot to its nearest neighbor and then the subjective contours (which can consist of several separate pieces) are established at those points where the sum of individual spread functions reaches a threshold, the peak value of the spread function. The consequence is that dots that are close to each other cluster into groups while those absent a close-enough neighbor become isolated. Once the subjective contours are established, the combined filled area within them can be gauged and used as the basis for numerosity judgment.
Although it fails for some patterns, this model is partially successful at explaining some of the effects spatial arrangement can have on numerosity judgment, as exemplified in the regular-random illusion (Fig. 6A).

Figure 6. Area models of numerosity. (A) The CODE model. The regular-random illusion is explained in terms of different areas enclosed by contours (Vos et al., 1988). (B) The occupancy model. Each object is assumed to occupy a territory of its own with a fixed radius (Allik & Tuulmets, 1991).

Allik and Tuulmets (1991) also considered the total area occupied by the pattern as the index to numerosity, the occupancy index, but instead of mandating a grouping process, they set out with the area taken up by individual dots. According to the so-called occupancy model (Fig. 6B), each dot occupies a circular territory of a constant radius that can overlap with the territory of other dots or not depending on the distance between them. If the center-to-center distance between two dots is closer than double the radius, their individual territories overlap and accordingly their combined contribution to the overall area occupied by the pattern is reduced, which in turn reduces the judged numerosity. The model found even greater success than the CODE model in explaining a number of numerosity illusions induced by differing spatial arrangements. However, it
doesn’t specify the way in which the radius of occupancy should be set, which makes the model sometimes too arbitrary. Apparently, a fixed radius is not going to accommodate all stimulus patterns.

More recently, in an attempt to battle the view of numerosity as a direct sense, Dakin et al. (2011) devised a model in which contrast energy is used to derive numerosity. Two Laplacian-of-Gaussian filters tuned to high and low spatial frequency (SF), respectively, are employed to filter the rectified input image. The ratio of response from these two filters pooled across all image locations with multiplicative random noise is posited to be a surrogate estimate of density, based on which a relative density judgment can be made. To derive a numerosity estimate, however, an additional process of gauging the pattern area is required. To this end, the density estimate obtained above is multiplied by the response of the low frequency filter, the candidate measure for area, to which additional random noise is introduced. This final operation represents an attempt to recover the high frequency response but with much noise added. The logic behind this is that high SF content of an image correlates highly with numerosity, making contrast energy in this band an ideal proxy for numerosity. However, according to the authors, the visual system doesn’t seem to rely directly on this measurement to estimate numerosity but instead derives density as a first step. The model explains the observed bias that a larger patch appears to be denser than a smaller patch with identical physical density and that the similar bias occurs with numerosity but to a lesser extent. Morgan, Raphael, Tibber, and Dakin (2014) proposed a very similar but more flexible mechanism by which, depending on stimulus properties and specific task, response from relevant filters measuring contrast
energy is used as proxy for numerosity, with scaling for contrast and blur. It is also flexible enough to allow selectivity for orientation, spatial frequency, contrast polarity and so on, as is necessary in some numerosity tasks. To summarize, the spatial filter models build on well-understood mechanisms of visual processing and has therefore provided a plausible solution to numerosity judgment. However, since these models are constructed mainly to support relative numerosity judgment, it is difficult to relate the measures they rely heavily on to the representation of absolute numerosity as proposed in many other studies.
Chapter 2. Study Overview

Although it’s yet unclear whether numerosity can be sensed directly or should be inferred from surrogate measures such as density, there has been unequivocal evidence that numerosity perception is not invariant to characteristics of a visual scene. Stimulus factors that have been found to influence numerosity judgment include element size, density, combined area (Gebuis & Reynvoet, 2012a, 2012b; Hurewitz, Gelman, & Schnitzer, 2006) and the spatial arrangement of elements, such as the convex hull of the pattern (Gebuis & Reynvoet, 2012a, 2012b), open space (Sophian & Chu, 2008), the clustering of elements (Ginsburg, 1991) and the depth ordering of elements (Schutz, 2012; Tsirlin, Allison, & Wilcox, 2012).

Surprisingly, few studies have investigated the effect of element contrast on numerosity judgment. Indeed one study discovered that a patch of dots with lower luminance is judged to be more numerous than a patch of dots with higher luminance (Ross & Burr, 2010). However, this effect only occurs at so low a luminance level as to potentially engage the scotopic system. Another study looked at the effect of element visibility (synonymous with contrast as relevant to the current study) on subitizing capacity but found null effect (Palomares & Egeth, 2010). In the current study the possible effect of element contrast on approximate numerosity perception will be systematically investigated. Above all, contrast is a basic property of any visual image and whether it contributes to numerosity judgment is in itself an interesting question to ask. Although the experiments proposed here are not designed to address the direct vs. indirect controversy mentioned earlier, the specific way in which element contrast affects
numerosity perception, if at all, will certainly shed light on the mechanisms of numerosity computation and provide useful constraints on computational models of numerosity.

We speculate why so few studies have looked at the effect of element contrast in the first place, in particular within the approximate number system. This lack of interest may well reflect a general failure in spotting any effect of supra-threshold contrast on numerosity perception. One potential reason for failure is the configuration of stimuli used in numerosity studies. Despite the fact that in real world situations objects belonging to one set are usually intermixed with objects from another set, the vast majority of previous studies on numerosity have used stimuli that contain isolated sets of objects, whether they were designed to have subjects estimate the numerosity of a single set or discriminate the numerosity between two or more sets (but see Cordes, Goldstein, & Heller, 2014).

The current study builds on the informal observation that when elements of different contrast are intermixed there appears to be an effect of contrast on perceived numerosity. More specifically, when equal numbers of light gray disks and white disks are intermixed on a dark gray background, there appears to be more gray disks than white ones, hence a numerosity illusion (see Fig. 7A). This phenomenon is counterintuitive in the sense that white disks seem more readily countable, while the gray ones, due to their similarity to the background, seem to be more difficult to individuate and therefore presumably more prone to underestimation.
For this dissertation, a series of experiments were conducted to address several aspects of the above-mentioned numerosity illusion:

Firstly, attempts will be made to quantify the illusion using a standard psychophysical procedure, which provides the basis for further exploring its properties and mechanisms by measuring the effect of several manipulations.

Secondly, the conditions under which the illusion occurs or in a complementary fashion disappears will be explored. In particular, the role of segregation in spatial and featural domain will be investigated.

Thirdly, the nature of the interaction that gives rise to the illusion will be probed. Specifically, it is asked whether the interaction between different contrasts is unidirectional or bidirectional and if the latter whether it’s symmetrical or asymmetrical. This will be examined by measuring how the perceived numerosity of each disk set is affected by the presence of the other set.

Lastly, task specificity of the illusion will be examined by comparing results from both numerosity discrimination and numerosity estimation tasks. It is particularly interesting to see if the effect persists when no explicit comparison of different contrasts is involved, which may help localize the stage of processing at which the illusion occurs. If the illusion arises at the perceptual stage rather than reflecting a decision bias at the response stage, we would expect it to occur across tasks.
Chapter 3. General Methods

3.1 Subjects

Observers in all experiments were students enrolled in an introductory psychology course and participated for course credit. They all had normal or corrected-to-normal vision and were naïve as to the purpose of the experiments.

3.2 Stimuli

The stimuli in most experiments consisted of disks (0.5° in diameter) distributed over a square patch measuring 10° each side (Fig. 7A). The patch was divided into a 12*12 grid of cells. On each trial, a subset of the 144 cells was randomly selected to place the disks. To increase randomness, the actual location of each disk was jittered both horizontally and vertically as much as 0.2° about the center of its cell. The overlap between disks was minimal. Disks of two different contrasts were intermixed on a dark gray background with a luminance of 18.2cd/m². The luminance of the low contrast (gray) and the high contrast (white) disks was 26.1cd/m² and 75.1cd/m², respectively. Since the disks were small relative to the background, physical contrasts are defined in Weberian terms as DL/L, where DL is disk luminance above the background and L is background luminance, and were (26.1-18.2)/18.2 = 0.43 for the low contrast disks and (75.1-18.2)/18.2 = 3.13 for the high contrast disks, both well above threshold. The stimuli were presented on a Dell LCD monitor and viewed at a distance of 60cm for 1.5s. Eye movements were not tracked and subjects were allowed to freely scan the image.
Figure 7. Example stimulus and psychometric function. (A) An example stimulus in a discrimination task. The reference disk set (usually the white ones) had a fixed numerosity of 50 while the numerosity of the test disk set (usually the gray ones) varied in a range of 30-70. In this example, both disk sets had a numerosity of 50. (B) An example psychometric function plotting the probability of the gray disk set being chosen as the more numerous one against the physical numerosity of gray disks. The numerosity value corresponding to a 50% probability on the curve is the PSE, in this case, 44. That is, 44 gray disks match 50 white disks in perceived numerosity.

3.3 Psychometric Procedures

3.3.1 Numerosity Discrimination Task

In the numerosity discrimination task, subjects were asked to judge the relative numerosity of the gray vs. white disks. A constant stimulus method was employed. In most experiments, the white disks were used as reference and had a standard (constant) numerosity of 50. The numerosity of gray disks was varied in the range of 30-70, with an increment of 5, resulting in 9 levels of physical numerosity. Each physical numerosity was compared against the standard numerosity and repeated 10 times throughout the experiment. A different pattern was used for each repetition of the same physical
numerosity. Each trial began with a fixation cross presented at the center of screen for 1s, followed by the disk display for 1.5s. The disk display was then replaced by a blank screen, upon which subjects had to make a response. Subjects were asked to press one key when they perceived more gray disks in the disk display than white ones and pressed another key for the opposite. The next trial started 3s after subjects had responded. For data analysis, the percentage of gray disks being selected as the more numerous was plotted against physical numerosity of gray disks and the data fit with a Weibull function (Fig. 7B). In this study, the PSE (point of subjective equality), which was derived as the physical numerosity of gray disks that corresponds to a 50% probability on the Weibull curve, was used as the dependent variable in all experiments, unless otherwise noted. A PSE less than 50 indicates an overestimation of gray disks relative to white ones and a PSE greater than 50 indicates the opposite. The slope of the Weibull curve was only of interest to characterize stability of the responses across subjects and conditions, as we had no theoretical expectations concerning this variable.

3.3.2 Numerosity Matching Task

A variant of the discrimination task was also used (Fig. 8). In this task, instead of using a single patch of disks presented at the center of the screen, two patches of disks were presented, one to the left and the other to the right of the fixation point, with a distance of 1° from the innermost edge of each patch to the screen center. One patch consisted of both gray and white disks created in the same fashion as described above, except that the numerosity of both gray disks and white disks was fixed at 50. The other patch consisted of either gray or white disks and varied in numerosity. Subjects were asked to judge the
relative numerosity of disks with the same contrast between the two patches. Similarly, a PSE was derived that indicated a match in numerosity. This task is used to investigate how the numerosity of a disk set would be influenced by the presence of another disk set in an intermixed state. For narrative purpose, in this dissertation any similar discrimination task involving two disk patches presented side by side will be called a matching task.

Figure 8. Example stimulus in a matching task. The intermixed patch always contains 50 gray disks and 50 white disks. The homogenous patch contains either gray or white disks that varied in numerosity. Subjects discriminated the numerosity of disks with the same color between the two patches. For this example, subjects were instructed to choose the patch with more white disks.

3.3.3 Numerosity Estimation Task

In the numerosity estimation task, instead of a forced choice, subjects had to estimate the absolute numerosity of disks and manually input their estimation using the number keys. The stimuli and procedure were generally similar to those used in the discrimination task, except that subjects reported the numerosity of either gray disks or white disks in response to a verbal prompt. The relevant disk set for report varied in numerosity in the
range 20-80 in steps of 10, while the numerosity of the irrelevant set was fixed at 50. A power function was fit to the data to describe the relation between reported numerosity and physical numerosity, consistent with earlier results showing that perceived numerosity is a compressive power of physical numerosity (Krueger, 1972; Krueger, 1982).
Chapter 4. Quantifying the Numerosity Illusion

The numerosity illusion that is the focus of this dissertation is that, when equal numbers of gray disks and white disks were intermixed, there appear to be more gray disks than white ones. In this part, attempts were made to quantify the illusion in the numerosity discrimination task. As just explained (in Chapter 3), one disk set served as the reference and had a fixed numerosity, while the numerosity of the other disk set was varied and compared to the reference numerosity. If the illusion is systematic, we would expect to obtain steep psychometrical curves and the PSEs would indicate that fewer gray disks than white ones are required for them to match in perceived numerosity. In Experiment 1, either white disks or gray disks were designated as the reference set and the other as the test. By comparing results of this manipulation within the same group of subjects, robustness of the numerosity illusion can be examined with respect to variation in procedure. In Experiment 2, we created a new display in which black disks and dark gray disks were intermixed on a light gray field, so that both disk sets had negative contrast, to discover that if a similar effect exists for this disk display as that in Experiment 1 where both disk sets had positive contrast.

4.1 Experiment 1: Gray vs. White Reference

As a first attempt to quantify the illusion, the discrimination task as described in the General Methods section was employed to measure how many gray disks would match 50 white disks in perceived numerosity. The PSE was predicted to be somewhere below 50, to reflect the observation that gray disks appeared to be more numerous than white ones when there were equal numbers of them. However, there is one caveat for this
procedure. Since the numerosity of white disks, which served as the reference, was fixed at 50 throughout, numerosity adaptation may be involved, which reduces the perceived numerosity of white disks and thus potentially accounts for the result. This is quite unlikely, though, given the fact that most naïve observers, when first presented an image depicting the illusion, were able to appreciate it immediately. However, numerosity adaptation may still play a role and inflate the illusion. To assess this possibility, in separate blocks, the gray disks served as the reference and had a standard numerosity of 50 while the white disks varied in numerosity on each trial. If numerosity adaptation plays a major role, we would predict that when gray disks were used as the reference, less than 50 white disks would be required to match 50 gray disks in perceived numerosity. If otherwise, more than 50 white disks would be required to achieve the match, to be consistent with the phenomenological observation. Moreover, since we adopted a within-subject design, we should expect a strong correlation between these two measurements of the same illusion when taken across individual subjects.

4.1.1 Methods
The stimuli and procedure as described in the General Methods section for the numerosity discrimination task were used, except that in separate blocks either white disks or gray disks served as the reference set. There were 90 trials for each reference type, which were divided evenly into two blocks with a short break in between. The test numerosity on each trial was randomly selected from the 9 numerosity levels, 30 through 70 in steps of 5. Therefore, there were 10 repetitions for each numerosity, based on which individual psychometric curves were derived. Eight subjects participated in the
experiment. Half of them completed the white reference blocks first and then the gray reference blocks while the other half received the opposite order. For both reference types, they performed the same forced-choice numerosity discrimination task. They were asked to press the left arrow key when they perceived more gray disks than white ones and press the right arrow key when they perceived the opposite, so the manipulation of reference set was transparent to the subjects. Before the experiment, subjects completed 10 trials of practice. The experiment lasted about 30 minutes.

Figure 9. Result of Experiment 1: white vs. gray reference. (A) Average PSEs (n=8). The dashed line indicates the standard numerosity. On average, 45 gray disks matched 50 white reference disks while 59 white disks matched 50 gray reference disks. Thus in both cases, numerosity of gray disks were overestimated relative to white ones. Error bar: ±1SE. (B) Individual PSEs. PSE for white reference measurement is along the x-axis and PSE for gray reference is along the y-axis. The solid line is the best-fitting linear regression line.

4.1.2 Results and Discussion

Psychometric curves were reliably steep for all of our subjects. For both reference types, average PSEs across subjects were significantly different from the reference numerosity
of 50 (Fig. 9A), \( t(7)=3.46, p=.01 \), and \( t(7)=4.50, p<.01 \), for white reference and gray reference, respectively. On average, 45 gray disks matched 50 white disks in perceived numerosity, amounting to a 10% illusion. More importantly, the illusion holds independent of which disk set served as the reference: an average of 59 white disks was required to match 50 gray disks in perceived numerosity, so in both cases, the numerosity of gray disks were overestimated relative to the white ones.

Inspecting individual data (Fig. 9B), numerically, except for one subject whose PSEs in both conditions were essentially 50, all had PSEs less than 50 in the white reference condition and PSEs greater than 50 in the gray reference condition, thus showing the same illusion of perceiving more gray disks than white ones in both conditions. A strong inverse correlation was obtained between PSEs from the two reference conditions, \( r=\cdot.85, p<.01 \), so that subjects who experienced a strong (or weak) numerosity illusion in one condition also did so in the other condition. Therefore, the illusion remained stable within individuals. These results cannot be explained by numerosity adaptation, according to which opposite effects should exist for the two reference conditions involved here. We conclude that, the role of numerosity adaptation, if any, seems to be minimal. For this reason, in all other experiments, if applicable, we used white disks as the reference and assumed that numerosity adaptation plays no major role.

4.2 Experiment 2: Negative Contrast

In Experiment 1, both disk sets had positive contrast. Although the term contrast was used, the illusion can equally be interpreted as an effect of luminance, since on a dark
field those disks with higher contrast also had a higher luminance level, and vice versa. So far the effect we observed is qualitatively similar to that reported by Ross and Burr (2010) who discovered that stimuli with a very low luminance were perceived to be more numerous than those with higher luminance. To disentangle the effect of contrast from that of luminance, in this experiment disks with negative contrast were used as stimuli in the numerosity discrimination task. If the illusion measured in Experiment 1 is a luminance effect, we would predict black disks to appear more numerous than dark gray disks when equal numbers of them were intermixed on a light gray field. If the illusion is instead a contrast effect, we would expect the opposite.

4.2.1 Methods

As illustrated in Fig. 10A, the disk display consisted of black disks (2.0cd/m²) and dark gray disks (8.2cd/m²) intermixed on a light gray field (35.4cd/m²). Black disks served as
the reference and had a standard numerosity of 50 while the numerosity of gray disks varied on each trial. The procedure was identical to that used in Experiment 1, except that there was only one reference condition in this experiment. Nine subjects participated and completed the numerosity discrimination task. They were instructed to press the left arrow key for perceiving more gray disks than black ones and press the right arrow key for the opposite. The experiment lasted about 15 minutes.

4.2.2 Results and Discussion

All but one subject had PSEs below 50 (Fig. 10B), so that the numerosity of gray disks were overestimated relative to black ones, indicating an illusion similar to that observed with positive contrast. The average PSE of all subjects (n=9) indicated that 44 gray disks matched 50 black disks in perceived numerosity, t(8)=3.64, p<.01, amounting to an illusion of 12%, a magnitude quite similar to that measured in Experiment 1. Once again the disks with lower contrast (but this time higher luminance) appeared to be more numerous. The combined results of Experiment 1 and 2 suggest that the numerosity illusion is indeed an effect of stimulus contrast rather than an effect of luminance.
Chapter 5. The Role of Segregation

Both Experiment 1 and 2 have clearly showed that numerosity judgment is affected by element contrast. Low contrast elements appeared to be more numerous than high contrast ones. However, since the illusion was observed under the special circumstance where the two disk sets were intermixed, it is yet unclear whether a contrast difference itself (without intermixing) is sufficient to induce the effect. To address this question, this part focuses on the role of segregation in the numerosity illusion. In Experiment 3, two disk sets like those used in Experiment 1 were segregated into different spatial regions in a numerosity discrimination task. We were interested in knowing whether the contrast-based numerosity illusion would persist even when the two sets of disks are physically segregated. In Experiment 4 and 5, instead of looking at physical segregation, the effect of perceptual segregation achieved by stereopsis and motion was probed. In Experiment 6, it was further asked how featural segregation due to orientation difference would affect the illusion. To investigate featural segregation, oriented bars were used as stimuli.

5.1 Experiment 3: Segregation by Visual Field

In this experiment, we repeated Experiment 1 using the identical procedure. However, the white and gray disks were either intermixed in one patch or segregated into two patches occupying separate spatial regions.

5.1.1 Methods

Stimuli and procedure in the ‘intermixed’ condition were identical to that used in Experiment 1, except that on different trials gray disks took on one of two gray levels:
26.1 cd/m$^2$, which was used in Experiment 1, and 45.8 cd/m$^2$. These are denoted as the low and high contrast conditions and were run to test the generalizability of the illusion over contrast levels. The luminance of white disks remained 75.1 cd/m$^2$ and the dark gray background 18.2 cd/m$^2$, so all disks had positive contrast. The white disks served as the reference set and had a standard numerosity of 50, while gray disks were the test set and varied in numerosity. In the numerosity discrimination task, subjects were asked to press the left arrow key when perceiving more gray disks (either gray level) and press the right arrow key when perceiving more white ones. For the ‘segregated’ condition, gray disks and white disks were segregated into different visual fields, forming two disk patches, one to the left and the other to the right of fixation (Fig. 11A). Each disk patch was created in the same fashion as the intermixed patch and placed with its innermost side at 1° eccentricity from the fixation. The side on which to present each disk set was balanced. Subjects were asked to pick the side with more disks by pressing the corresponding left or right arrow key. All other aspects were identical to the ‘intermixed’ condition. Six subjects participated in this experiment and completed the ‘intermixed’ and ‘segregated’ conditions in separate blocks. Half of them completed the ‘intermixed’ blocks first and half of them completed the ‘segregated’ blocks first. The experiment lasted about 50 minutes.

5.1.2 Results and Discussion

PSEs were derived from individual psychometric curves and the average of all subjects (n=6) were shown for different conditions in Fig. 11B. The numerosity illusion was replicated for both gray levels in the ‘intermixed’ condition, $t(5)=3.90$, $p=.01$, and
Figure 11. Example stimulus and result of Exp. 3: segregation by visual field. (A) An example stimulus in the segregated condition. The two disk sets were segregated to the left and right of fixation. The white disk set had a fixed numerosity of 50 while the numerosity of gray disks varied. In this example, there are 50 gray disks. The gray disks took on one of two gray levels: 26.1 cd/m$^2$ (low) and 45.8 cd/m$^2$ (high). (B) Mean PSEs plotted for all conditions. Error bar: ±1SE.

$t(5)=3.64$, $p<.05$, for low and high gray level, respectively, thus revealing a robust effect. On average, 43 low-contrast gray disks and 41 high-contrast gray disks were required to match 50 white disks in perceived numerosity, but this numerical difference in illusion
strength was not significant, t(5)=0.56, p=.60, suggesting that the illusion is relatively insensitive to the amount of contrast difference between the reference and test disks. When the two disk sets were segregated into different visual fields, however, the numerosity illusion was eliminated altogether. The PSEs for both gray levels (51.1 and 51.4, for low and high, respectively) were not significantly different from the standard numerosity of 50, t(5)=2.03, p=.10 and t(5)=2.00, p=.10, for low and high, respectively. Therefore, subjects showed veridical comparative judgment of numerosity in the ‘segregated’ condition. If anything, there was a slight trend for subjects to overestimate the numerosity of white disks relative to gray ones when they were segregated into two separate patches, opposite to the illusion observed when they were intermixed. These results suggest that the mere presence of a contrast difference is insufficient to induce the numerosity illusion and that intermixing constitutes a necessary condition for the contrast-based illusion to occur.

One question remains: provided a contrast difference between two sets of disks, is intermixing sufficient to induce the numerosity illusion? Can other forms of segregation break the illusion? The construction of random dot stereogram and kinematogram has convincingly demonstrated that stereopsis and motion cues strongly contribute to the segregation of image elements occupying the same retinal region, giving rise to a vivid impression of transparency (Julesz, 1971). In the following experiments we ask if perceptual segregation based on stereopsis and motion grouping would also destroy the numerosity illusion just as physical segregation did, even though in these cases the two sets of disks are still intermixed retinally. We will eventually go further to see whether
the numerosity illusion can be replicated using oriented bars instead of disks and whether segregation based on orientation impairs the illusion.

5.2 Experiment 4: Segregation by Stereoscopic Depth

In this experiment, stereopsis was used to create a depth structure within the disks. The purpose is to segregate gray disks and white disks into different depth layers, so that the two disk sets are perceptually isolated. If segregation in general breaks the numerosity illusion, we would expect the two disk sets to show a veridical match in perceived numerosity when placed at different depth layers. However, depth manipulation may have some unintended consequences that may complicate the interpretation of any results that come about. As one possibility, depth may trigger the size constancy mechanism, so that disks farther away from the observer increase in perceived size relative to disks that are closer, which in turn may affect numerosity judgment. Since this experiment is not designed to address the effect of depth per se and have no specific predictions in this regard, we tried to circumvent this complication by creating a relatively small depth separation between different layers so as to minimize the change in perceived disk size with distance. We also included two depth orders: ‘white in front’ and ‘gray in front’. The hypothesis is that segregation by depth eliminates the numerosity illusion, regardless of what depth order is there between the two disk sets. A ‘flat’ condition was also included, in which the two disk sets lay on the same depth plane, a replication of Experiment 1 but now implemented with a stereoscopic setup.

5.2.1 Methods
As illustrated in Fig. 12A, the disk display consisted of two images, each presented to either the left or right eye with the help of a stereoscope. A positive disparity of 2 arcmin was added to the gray disk set to push them behind the white disks, i.e., the ‘white in front’ condition, and a negative disparity of 2 arcmin was added to bring them forward, i.e., the ‘gray in front’ condition. No disparity was applied for the ‘flat’ condition. The three depth conditions were randomized throughout the experiment. To allow adequate time for subjects to stabilize the image and see depth, the disk display was presented for 2s, a duration that was 0.5s longer than used in other experiments. Subjects completed a numerosity discrimination task with all other aspects identical to that of Experiment 1. Since a relatively small binocular disparity was used, in order to make sure that our subjects saw depth properly, before the numerosity task, a pretest was administered in which subjects were asked to report the depth order of the disk sets: ‘white in font’, ‘gray in front’ or ‘flat’. The pretest comprised 12 trials for each depth condition and represented a subset of the stimuli used for the numerosity task, spanning the whole range of test numerosities. Only those subjects who passed the pretest with accuracy over 90% were included. Ten subjects participated in this experiment, but two of them failed the pretest and were excluded from data analysis. The experiment lasted about 45 minutes.

5.2.2 Results and Discussion

Despite the difference in experimental setup from previous experiments, the numerosity illusion was nevertheless replicated in the ‘flat’ condition (Fig. 12B). On average, 46 gray disks matched 50 white disks in perceived numerosity and the difference is significant, t(7)=4.44, p<.01. By contrast, average PSEs for both ‘white in front’ and ‘gray in front’
Figure 12. Example stimulus and result of Exp. 4: depth segregation. (A) Simulation of the stereoscopic stimulus used in Experiment 5a. A small relative disparity is added. If cross-fused, the white disks are seen in front. In this example, there are 50 white and 50 gray disks. (B) Mean PSEs plotted for the depth conditions. Error bar: ±1SE.
segregation eliminates the contrast-based numerosity illusion, regardless of the specific depth order. Since there is clearly no difference between the ‘white in front’ and ‘gray in front’ conditions, it can be concluded that segregation rather than depth order was critical.

These results are interesting for two reasons. Firstly, they revealed that the contrast-based numerosity illusion is susceptible not only to physical segregation as demonstrated in Experiment 3 but also to perceptual segregation due to stereopsis, despite the fact that in the latter case the two disk sets are still retinally intermixed. This suggests that a segregation process is attempted before numerosity is extracted from a visual image and that the numerosity illusion we have repeatedly observed in this study results at least partly from an incomplete segregation between the two disk sets. However, why a failure at segregation should lead to the illusion is itself an interesting question to ask and will be discussed in other parts of this dissertation. Secondly, these results also have implications for a potential explanation of the numerosity illusion. According to this explanation, in our stimulus patch of intermixed gray and white disks, contrast serves as a depth cue so that white disks are seen in front of gray ones. Our visual system implicitly assumes that some gray disks are occluded by white disks and as a consequence automatically compensates for the numerosity of gray disks. To the extent that contrast is an effective depth cue and the depth impression thus established is equivalent to that by stereopsis, the results of this experiment clearly refute this hypothesis. If the occlusion hypothesis were true, we would predict that placing the gray disks at back should strengthen the numerosity illusion while placing them in front would counteract or even reverse the illusion. However, this is clearly not what happened.
5.3 Experiment 5: Segregation by Motion

Experiment 4 demonstrated that segregation by depth eliminated the numerosity illusion. Motion is another effective cue that facilitates image segregation. If segregation in general destroys the numerosity illusion, we would expect the same to happen with segregation by motion as that with segregation by depth. This, if true, would strengthen the notion that there exists a mandatory segregation process before the computation of numerosity. When segregation fails, numerosity judgment may suffer. In this experiment, we set out to test the hypothesis that the contrast-based numerosity illusion is also susceptible to segregation by motion. To this end, the two disk sets were put in anti-phase sinusoidal motion (or as we call it the ‘out-of-phase’ motion condition) so that each disk set coheres by itself and contrasts in motion direction with the other set. This creates a vivid impression of segregation. As our main interest here is in motion-induced segregation, we also introduced the ‘in-phase’ motion condition, in which the two disk sets were set in common motion and thus cohere as a whole, for the purpose of ruling out any effects motion per se may have on perceived numerosity. One additional condition, the ‘static’ condition, was simply a replication of what was done in Experiment 1 and served as the benchmark condition.

5.3.1 Methods

The stimuli and procedure were identical to that of Experiment 1, except that motion was introduced to the disks (Fig. 13A). In the ‘in-phase’ motion condition, both disk sets assumed the identical sinusoidal motion as a whole with a frequency of 2Hz and an amplitude of 0.2°. In the ‘out-phase’ motion condition, the same sinusoidal motion was
Figure 13. Example stimulus and result of Exp. 5: motion segregation. (A) Motion stimulus used in Experiment 5b. Colored arrows are added for illustrative purpose only, not actually shown in the experiment. All disks participated in a sinusoidal motion. (B) Mean PSEs plotted for the motion conditions. Error bar: ±1SE.

applied but with a phase difference of 180° between the two disk sets so that they always moved in opposite directions. In the ‘static’ condition, all disks were static, just as they were in Experiment 1. The three conditions were randomized throughout the experiment.
Eight subjects participated in the experiment and performed the numerosity discrimination task. The experiment lasted about 45 minutes.

5.3.2 Results and Discussion

As shown in Fig. 13B, the numerosity illusion was again replicated in the ‘static’ condition, with on average 44 gray disks to match 50 white disks in perceived numerosity, t(7)=4.43, p<.01. The illusion was also evident for the ‘in-phase’ motion condition, with an average PSE of 43 gray disks, which is significantly different from the reference numerosity, t(7)=5.15, p<.01. There was clearly no difference in illusion strength between these two conditions, suggesting that motion alone has little effect on the illusion. However, the numerosity illusion disappeared for the ‘out-phase’ motion condition, as the average PSE of 49 is hardly different from the reference numerosity of 50, t(7)=1.26, p=.25. This indicates that segregation by motion is sufficient to block the numerosity illusion, just as effectively as segregation by depth.

Results of this experiment corroborate the notion that segregation of the two sets of disks from each other destroys the contrast-based numerosity illusion. It’s worth noting that, however, by segregation, we mean the segregation between the two disk sets rather than any other forms of segregation. Presumably, the ‘in-phase’ motion introduced in this experiment has facilitated the segregation of both disk sets from the background. However, segregation in this sense has little effect on the numerosity illusion. Conceivably, it is the segregation between the two disk sets that is the critical step towards accurate numerosity judgment. On a separate note, it is perhaps not too
surprising for segregation by motion to have the same effect as segregation by stereopsis, since there is good evidence that contrasting motion leads to perception of depth separation (Snowden & Verstraten, 1999), so that segregation by motion may be just another instantiation of segregation by depth.

5.4 Experiment 6: Segregation by Orientation

Our purpose for this experiment is two-fold: to replicate the numerosity illusion with stimuli that are drastically different in appearance from that of disks; and to explore the effect of featural segregation on the illusion. To this end, we substitute oriented bars, which could either be horizontal or vertical, for the disks used in previous experiments. It is predicted that the contrast-based numerosity illusion would not be confined to disks but also generalize to oriented bars as the target of enumeration. However, the effect would be dependent on whether the two sets of bars share orientation or not. When segregation is achieved due to orientation contrast, we expect the illusion to diminish or even disappear. To test this hypothesis, we set up a condition in which all gray and white bars had the same orientation, either horizontal or vertical, i.e., the ‘same’ condition, and one in which all the gray bars shared one orientation that was orthogonal to the orientation shared by all the white bars, i.e., the ‘orthogonal’ condition. We also included an intermediate condition, in which half of each bar set were horizontal and the other half were vertical, the so-called ‘half-half’ condition. It is hypothesized that this condition represents an intermediate state in terms of segregation, halfway between the ‘same’ and ‘orthogonal’ conditions; accordingly, a numerosity illusion of an intermediate magnitude should be expected for this condition.
Unfortunately, it has been reported that, when subjects were asked to judge the proportion of each orientation in an intermixed patch of horizontal and vertical bars, they overestimated the proportion of horizontal bars relative to vertical bars (Shuford, 1961), which, if replicable, inevitably adds a layer of complication to the current experiment. Taking this into account, we decided not to pool across orientations but instead, for the ‘same’ and ‘orthogonal’ conditions mentioned above, we also looked at specific orientations. Since subjects were not able to complete all conditions in one hour, we ended up running two sub-experiments, one in which we included only the ‘same’ and ‘half-half’ conditions and the other in which we included only the ‘orthogonal’ and ‘half-half’ conditions.

5.4.1 Experiment 6a

Methods

As shown in Fig. 14A, oriented bars, 0.6° in length and 0.2° in width, which could either be horizontal or vertical, were used as the elements of the display. There were gray bars and white bars, which had the corresponding luminance of the gray and white disks used in previous experiments, intermixed on the same gray field. However, because of the unique dimension of oriented bars, to avoid individual bars touching, the standard 10° by 10° square patch was divided into an imaginary grid of 10*10 cells and a bar was placed at the center of each cell without position jittering. Therefore, there were always 100 bars in each display, with the numerosity of gray bars and white bars covarying in opposite direction. Accordingly, there was no reference numerosity and the psychophysical procedure was aimed at locating the point where N gray bars would match 100-N white
bars. Using the method of constant stimuli, the numerosity of gray bars was varied in the range of 40-60 by steps of 2 and the numerosity of white bars varied accordingly. Psychometric curves were again based on 10 repetitions of each numerosity level. Note the change in meaning of PSE in this experiment. For example, if PSE equals 48, this means that 48 gray bars match 100-48=52 white bars in perceived numerosity, giving a difference of 4 (52-48) rather than 2 (50-48), so PSEs in this experiment will be twice as close to 50 as in previous experiments even when the illusion is of a comparable strength. All other aspects of this experiment were similar to that of previous experiments.

This experiment included the ‘same’ and the ‘half-half’ conditions as described above. For the reason mentioned earlier, the ‘same’ condition in turn included the ‘both horizontal’ and ‘both vertical’ sub-conditions, which were treated as separate conditions when measuring the PSEs, so effectively there were three conditions with equal number of trials for each. All these conditions were randomized throughout the experiment. Subjects performed a numerosity discrimination task and were asked to press the left arrow key when perceiving more gray bars than white ones and press the right arrow key when perceiving the opposite. Eight subjects participated in this experiment. The experiment lasted about 50 minutes.

Results
One subject who was a outlier for all three conditions (two standard deviations off the group mean) was excluded from data analysis. Exclusion of this subject did not cause major changes to statistics, with one exception to be noted below. For better comparison
with Experiment 6b, we chose to present here the results without this subject (Fig. 14B). Average PSEs of ‘both horizontal’ and ‘both vertical’ conditions were 46 and 45, respectively, both significantly different from 50, t(6)=4.52, p<.01, and t(6)=4.59, p<.01, respectively. Note that although no reference numerosity was used in this experiment, 50 represents the point of veridical match between gray and white bars and thus is the expected value to compare actual PSEs against. Therefore, when gray bars and white bars shared a common orientation, the numerosity of gray bars was overestimated relative to the white bars, a numerosity illusion similar to that observed in other experiments with disks as stimuli. For the ‘half-half’ condition, the PSE was 48, also significantly different from 50, t(6)=3.58, p=.01. However, if the outlier was included, although the trend remained the same, the difference was no longer significant. This uncertainty regarding the presence or absence of a numerosity illusion for the ‘half-half’ condition will be resolved in Experiment 6b.

It is hypothesized that the illusion in the ‘half-half’ condition, if occurring at all, should be weaker than in the ‘same’ orientation condition. To test this prediction, since ‘both horizontal’ and ‘both vertical’ conditions have given quite similar results, the PSEs from these two conditions were averaged as the ‘same’ condition and compared to that of the ‘half-half’ condition. Indeed, a significant difference was found, t(6)=2.43, p=.05. Since local segregation due to orientation contrast may exist between the gray and white bars in the ‘half-half’ condition, this result fits well with the notion that the numerosity illusion depends inversely on the degree of segregation between the two sets.
Figure 14. Example stimuli and results of Exp. 6: segregation by orientation. (A) Oriented bar stimuli. Same: white and gray bars have the same orientation; Orthogonal: white and gray bars have orthogonal orientations; Half & Half: half of each set have horizontal and the other half vertical orientation. In these examples, both white and gray bars have a numerosity of 50. Exp. 6a included the ‘same’ and the ‘half-half’ conditions; Exp. 6b included the ‘orthogonal’ and the ‘half-half’ conditions. (B) Mean PSEs plotted for the orientation conditions in Exp. 6a. (C) Mean PSEs plotted for the orientation conditions in Exp. 6b. Error bar: ±1SE.

5.4.2 Experiment 6b

Methods

This experiment included the ‘orthogonal’ and the ‘half-half’ conditions. For the same reason, as in Experiment 6a, the two orientation combinations that belong to the ‘orthogonal’ condition were treated as separate conditions: ‘white horizontal & gray vertical’ and ‘white vertical & gray horizontal’, for simplicity, abbreviated as the ‘HV’ and ‘VH’, respectively, with the first letter standing for the orientation of white bars and the second for gray bars. Apart from this difference in experimental conditions, all aspects of stimuli and procedure were identical to that of Experiment 6a. Eight subjects participated in this experiment.

Results

As shown in Fig. 14C, PSEs for both ‘VH’ and ‘HV’ conditions were about 51; neither were significantly different from 50, t(7)=1.68, p=.14, and t(7)=1.60, p=.15, respectively, thus indicating no numerosity illusion for both conditions. There was no evidence for a bias to overestimate the numerosity of horizontal bars (Shuford, 1961), as this would predict different results for the ‘VH’ and ‘HV’ conditions. Critically, for the ‘half-half”
condition, a numerosity illusion was again observed, with an average PSE of 48, which was significantly different from 50, \( t(7)=2.98, p=.02 \). PSEs of the ‘VH’ and ‘HV’ conditions were then pooled as the ‘orthogonal’ condition and compared to the ‘half-half’ condition; a significant difference was found, \( t(7)=5.81, p<.01 \). Therefore, when the two sets of bars were segregable by orientation, the numerosity illusion was eliminated and subjects were able to make veridical numerosity judgment. However, when segregation was incomplete, as in the ‘half-half” condition, the illusion remained.

5.4.3 Discussion

Putting Experiments 6a and 6b together, a clear picture emerges. First of all, using oriented bars as stimuli, we confirmed the generality of the numerosity illusion previously observed with disks. When equal numbers of gray bars and white bars with the same orientation were intermixed on a dark gray field, there appear to be more gray bars than white ones. Therefore, elements that are lower in contrast tend to be higher in perceived numerosity than similar elements with a higher contrast. It is possible that the illusion also applies to more complex stimuli. For instance, if we create a mixture of cartoon faces from two sets differing only in luminance contrast, it is likely that the low-contrast faces should appear more numerous than the high-contrast ones. It is quite tempting to have this idea tested in the future, which if proven true, would implicate that the numerosity illusion not only operates at the feature level but also at the object level.

The use of oriented bars has also provided us with an opportunity to test the role of featural segregation in the numerosity illusion. It turned out the presence or absence of
the illusion depended on the orientations of the two sets of bars. The illusion occurred when the two sets shared the same orientation but not when they took orthogonal orientations. Previous experiments in this section demonstrated that segregation in general breaks the illusion. Experiment 3 showed that physically segregating the two sets of elements eliminated the numerosity illusion. Experiment 4 and 5 further showed that perceptual segregation established by stereopsis and motion grouping also destroyed the illusion. These experiments have lent support to the notion that image segregation is a mandatory stage in numerosity computation, the failure of which may cause bias in numerosity judgment. However, these experiments were all concerned with segregation in the spatial domain. The current experiment has advanced the idea by showing that the numerosity illusion is also susceptible to featural segregation. When two sets of bars were segregable based on orientation, they match veridically in perceived numerosity. It is also interesting that the effect of segregation doesn’t take place in an all-or-none fashion. When the two sets of bars were only partially segregable by orientation contrast, due to hypothesized local segregation, as was the case in our ‘half-half’ condition, a less strong but still significant numerosity illusion was observed, suggesting that segregation impacts numerosity perception in a graded fashion. As the degree of segregation increases, the bias in numerosity judgment decreases. An alternative interpretation of the results of the ‘half-half’ condition is that the image was first segregated based on orientation; after segregation, the numerosity illusion occurs on a smaller scale for each orientation since there were both gray bars and white bars with that particular orientation. One potential avenue to manipulate the degree of segregation and tease apart these two mechanisms is to systematically vary the orientation difference between the two sets of oriented bars and
meanwhile measure the numerosity illusion. Whatever specific mechanism was responsible in this condition, our results highlighted the important role of segregation in numerosity perception.

It should be acknowledged that contrast is itself a cue to segregation but apparently it is too weak a cue to lead to complete segregation (Beck, Graham, & Sutter, 1991). In the context of the current experiment, segregation based on orientation competes with segregation based on contrast when the two cues are in conflict. Since orientation selectivity is a well-known property of visual processing (De Valois, Yund, et al., 1982; Hubel & Wiesel, 1968), it is perhaps not surprising that orientation is the main factor that determines segregation. Following from this argument, gray elements and white elements were segregated to an even less degree when oriented bars with the same orientation were used as the stimuli than when circular elements like disks were used. This may potentially explain the observation that in the current experiment, the numerosity illusion appears to be greater in strength than in the previous experiment with disks as the stimuli. In the current experiments, 45 gray bars matched 55 white bars in perceived numerosity while the typical scenario in previous experiments was for 45 gray disks to match 50 white ones. However, besides the use of different stimuli, there were also other important methodological differences between these experiments, therefore this comparison of illusion magnitude is not straightforward.
Chapter 6. Revealing the Interaction: Matching Experiments

In previous sections, we have reported a numerosity illusion in which low-contrast (gray) elements tend to have higher apparent numerosity than high-contrast (white) elements. It has been further demonstrated that the illusion occurs only when the two sets of elements are intermixed and not perceptually segregated. The question that remains is, when gray elements and white elements fail to segregate from each other, what causes the numerosity bias? The experiments in this section are aimed at uncovering the interactions between gray and white elements while they are in a mixture, in the condition that gives rise to the numerosity illusion. Specifically, we are interested in knowing how the presence of one element set affects the perceived numerosity of the other set. Cordes and colleagues (Cordes et al., 2014) have reported a contextual effect in which numerosity estimation of one set of objects was decreased when accompanied by another set in an intermixed patch. If a similar contextual effect is responsible for the numerosity illusion discovered in this study, an asymmetry should be assumed for the mutual influence between gray and white elements, so that the presence of gray elements reduces the perceived numerosity of white elements to a larger extent than the other way around, to explain the results already presented in this dissertation.

Another obvious explanation for the numerosity illusion depends upon the concept of misclassification. Since our targets were defined in terms of lightness, which is a continuous property, it is possible that some gray elements are misclassified as white ones and some white elements misclassified as gray ones, with the latter happening more frequently than the former. Yet another not so obvious explanation has to do with
occlusion. According to this explanation, contrast serves as a depth cue so that white elements are seen in front while gray elements recede to the back. Due to this depth arrangement, the visual system implicitly assumes that some gray elements are occluded by white elements and in an effort to recover this loss compensates for the numerosity of gray elements. This account of the numerosity illusion by unconscious inference is quite unlikely, in light of the results of Experiment 5 in which the depth order of the two sets of elements were manipulated. The illusion disappeared no matter which set was in front.

To reveal the interactions between the two sets of elements, all experiments in this section employed the matching procedure as described in the General Methods section. Instead of judging the relative numerosity of gray vs. white elements within a single patch, subjects were asked to make numerosity judgments across two patches. Using this procedure, it is feasible to make measurement selectively on one set of elements from the intermixed patch in comparison to an uncontaminated set in another patch, thus possibly shedding light on the inner workings of the illusion. Experiment 7 was designed to test the misclassification theory. Experiment 8 was aimed at revealing the nature of mutual influence between the two element sets. Experiment 9 was designed to test the perceptual suppression account we advocate to explain the numerosity illusion.

6.1 Experiment 7: Superset Matching

According to the misclassification hypothesis, gray elements can be misclassified as white elements and white elements can be misclassified as gray. However, since each individual element is well above threshold and clearly visible, misclassification among
the elements should not have any impact on the total numerosity of the superset including all the elements regardless of their contrast. One straightforward way to test this prediction is to match this intermixed patch in terms of numerosity to a second patch that is not contaminated by the numerosity illusion in question and can be used as a benchmark.

6.1.1 Methods
The stimuli and procedure were similar to that described in the General Methods section for a matching task (Fig. 2). The intermixed patch consisted of 50 white disks and 50 gray disks throughout the experiment. The other patch consisted of disks of a homogenous contrast and varied in numerosity in the range of 80-120 by steps of 5. The luminance of disks in this homogenous patch was 45.8cd/m², approximately halfway between the luminance of the gray and white disks in the intermixed patch. This luminance value was chosen to approximately equate the total luminance of the two patches while avoid potential response bias that may arise if either gray or white disks were used. However, what specific luminance value to use does not seem to matter according to Experiment 3, the results of which showed that numerosity match between homogenous patches doesn’t depend on the luminance levels of the disks used. Subjects performed the numerosity matching task and were asked to select the patch with more disks regardless of color (gray or white) by pressing the corresponding left or right arrow key. Eight subjects participated in this experiment, which lasted about 20 minutes.

6.1.2 Results and Discussion
The PSE in this experiment indicates the numerosity of the homogenous patch that matched the intermixed patch in perceived numerosity. As can be seen in Fig. 15, all subjects had a PSE below 100, which was the physical numerosity of the intermixed patch. This means that all subjects underestimated the numerosity of the intermixed patch relative to the homogenous patch. The average PSE was 91, which was significantly different from 100, $t(7)=3.29$, $p=.01$. This result was clearly at odds with the prediction of the misclassification hypothesis, according to which the total numerosity of the intermixed patch should be preserved despite the numerosity illusion arising within that patch. Apparently, some disks were missing in the intermixed patch and misclassification alone cannot explain why the numerosity illusion occurs.

![Figure 15. Result of Exp. 7: superset matching. Individual PSEs are plotted (n=8).](image)

However, there is one explanation that may salvage the misclassification hypothesis. It is possible that the total numerosity of the intermixed patch cannot be perceived directly but is added up from the numerosity of each of the two sets it contained. In other words, the visual system automatically parses the intermixed patch into a gray set and a white set,
with the numerosity mechanism operating over each set. The cost associated with parsing and adding numerosity up explains the numerosity underestimation of the intermixed patch. By this account, misclassification between gray and white disks may happen during the parsing process, with more white disks misclassified as gray than the opposite. However, it has been reported by other researchers that subjects tend to overestimate the numerosity of a heterogeneous array of colored disks relative to a homogenous one (Cordes et al., 2014). Therefore, although the parse-and-add account seems appealing, it is not a parsimonious one and more importantly lacks empirical support.

Moreover, the results of this experiment were also not compatible with the occlusion hypothesis, which would predict an increase in total numerosity of the intermixed patch. We speculate that, instead of misclassification or occlusion, some kind of perceptual suppression takes place within the intermixed patch, which provides a better account for both the numerosity illusion and the results of this experiment. This new account will be explored in the following experiments.

6.2 Experiment 8: Gray/White in Intermixture vs. G/W in Isolation

Experiment 7 demonstrated that some disks are ‘missing’ in the intermixed patch presumably due to the interaction between the two disk sets. Since the numerosity illusion is that there appear to be more gray disks than white ones, it should necessarily follow that white disks are ‘missing’ more than gray disks. To test this prediction, in this experiment, the perceived numerosity of either white disks or gray disks in the intermixed patch was measured by selectively matching each set to a set of disks presented in
isolation in a second patch, to reveal how intermixing differentially affects the perceived numerosities of the gray and white disks.

6.2.1 Methods

The stimuli and procedure were described in the General Methods section, see Fig. 2. Unlike Experiment 7, in this experiment, subjects made numerosity judgment based on color. The homogenous patch consisted of either white or gray disks that varied in numerosity in the range of 30-70 in steps of 5. Subjects were asked to select the patch with more disks of the same color, either gray or white, contingent on which color was in the homogenous patch. The PSE would indicate how many disks in the homogenous patch would match 50 disks of the same color in the intermixed patch, so the numerosity of the homogenous patch serves as a benchmark against which to measure the perceived numerosity of each set in the intermixed patch. Six subjects participated in this experiment, which lasted about 30 minutes.

Figure 16. Result of Exp. 8: intermixture vs. isolation. Mean PSEs are plotted. Error bar: ±1SE.
6.2.2 Results and Discussion

The results were shown in Fig. 16. On average, 42 white disks in isolation matched 50 white disks in intermixture, the difference being significant, \( t(5)=5.98, p<.01 \); 51 gray disks in isolation matched 50 gray disks in intermixture, with no significant difference, \( t(5)=1.40, p=.22 \). Therefore, white disks decreased in perceived numerosity when intermixed with gray disks while the perceived numerosity of gray disks remained unaffected by the presence of white disks. This result is both qualitatively and quantitatively consistent with the numerosity illusion we measured before: approximately 40-45 gray disks were required to match 50 white disks in perceived numerosity in an intermixed patch. This result is also consistent with the result of Experiment 7 and indicates that the decrease in perceived numerosity of the superset is due exclusively to the ‘missing’ of white disks.

The fact that white disks are missing to a large extent is counterintuitive in the sense that white disks are readily visible in both patches so that a veridical match should be expected. On the other hand, the veridical perception of the numerosity of gray disks speaks against any theories that would predict a change, either an increase or decrease, in the perceived numerosity of this set. Specifically, both the misclassification and occlusion hypotheses predict that the perceived numerosity of gray disks should increase in the intermixed patch, although for different reasons. To account for this asymmetry in the interaction between the two sets of disks, here we advocate a perceptual suppression hypothesis: when gray disks and white disks are intermixed, signals from the white disks that are used for numerosity computation are suppressed by the presence of nearby gray
disks. Moreover, the suppression is unidirectional in nature so that signals from the gray disks are not affected. Since how numerosity is derived from a visual image is itself a question in fierce debate (Anobile et al., 2014; Dakin et al., 2011), the account we are proposing to explain our numerosity results is necessarily vague on other matters. Nevertheless, the hypothesis will be further explored in Experiment 9, and more details regarding how the suppression mechanism operates will be developed in the General Discussion section.

6.3 Experiment 9: Matching in Noise

Both misclassification and occlusion hypotheses presuppose an individuation stage that should take place prior to any interaction between the gray and white disks. Only after each disk has been individuated is there a ground for misclassification or occlusion to happen. By contrast, no such presupposition should be made for the perceptual suppression hypothesis, according to which suppression acts upon the raw signal used to derive numerosity, whatever form it may take. By this account, each disk set in the intermixed patch simply serves as visual noise to the other disk set for the purpose of numerosity perception. This idea is explored in the current experiment by substituting a cloud of texture-like elements for one set of the clearly individuated disks used in the previous experiment while measuring the perceived numerosity of the other set. It is hypothesized that the perceived numerosity of disks embedded in such noise should be affected in the same way as in an intermixed patch of white and gray disks, to the extent that the cloud and the disks share spatial frequency content.
Figure 17. Example stimulus and result of Exp. 9: matching in noise. (A) Noise stimulus. There are two patches of disks on each trial, onto one of which was added a cloud of tiny disks as noise. Gray (white) tiny disks were added to a patch of white (gray) disks. The noise patch had a fixed numerosity of 50 while the numerosity of the other patch varied within the range of 30-70. In this example, both patches contain 50 white disks. (B) Mean PSEs. Error bar: ±1SE.

6.3.1 Methods

The stimuli and procedure were identical to that used in Experiment 8, except that one set of disks in the intermixed patch, which was the irrelevant set for matching on each trial, was replaced by 300 tiny disks (0.2° in diameter) of the same contrast randomly distributed over the space between the large disks that were the targets for matching (Fig.
17A). Therefore, on some trials large white disks were embedded in gray texture-like noise while on the other trials large gray disks were embedded in white texture-like noise. The cloud of tiny disks had approximately the same aggregate area as the original set of large disks. On each trial subjects discriminated the numerosity of the large disks, either white or gray, between two patches, one with noise and the other without. Six subjects participated in this experiment, which lasted about 20 minutes.

6.3.2 Results and Discussion

As shown in Fig. 17B, the result pattern was quite similar to that of Experiment 8. On average, 39 white disks in isolation were required to match 50 white disks embedded in noise, the difference being significant, t(5)=8.64, p<.01; 47 gray disks in isolation were required to match 50 gray disks in noise, the difference being insignificant, t(5)=1.47, p=.20. So again, some white disks were ‘missing' while all gray disks remained. This result basically confirmed our hypothesis that the interaction between the two contrasts that gives rise to the numerosity illusion operates at a relatively early stage of visual processing. The interaction does not seem to rely on two sets of segmented objects and an explicit individuation process seems unnecessary for the numerosity illusion to occur.

This experiment has also provided additional opportunity to test both the misclassification and occlusion hypotheses. Since the large target disks and the tiny distractor disks constituting the noise were drastically different in appearance, both misclassification and occlusion (of the target disks) were unlikely to happen with the current stimulus configuration. Nevertheless, the same result was obtained as before: gray
disks won over white disks in perceived numerosity, suggesting that an account of the numerosity illusion that resorts to lower levels of processing is more plausible.
Chapter 7. Absolute Numerosity Estimation

In the previous sections, a numerosity illusion was reported, in which gray disks appear to be more numerous than an equal number of white disks when they are intermixed on a dark gray field. However, this illusion is vulnerable to segregation by various means. It has been further demonstrated that the illusion is due exclusively to a decrease in the perceived numerosity of white disks. While our results have been quite consistent despite a range of variations in stimulus configuration and procedures, all these experiments conducted so far employed a numerosity discrimination task in which subjects explicitly compare the numerosity of two sets of disks. This kind of relative judgment may be prone to bias induced by factors other than numerosity per se (Gebuis & Reynvoet, 2012a). It is questionable whether subjects rely on a numerosity representation to make discrimination at all. As one possibility, subjects may base their decisions on pattern comparison between the two sets of disks. It is also possible that the numerosity illusion we reported may reflect some kind of post-perceptual response bias. It should be noted that most of our subjects didn’t have any doubts that they were making numerosity judgments rather than anything else, but one may be unaware of biasing influences.

In this section, a numerosity estimation task, in which subjects had to estimate the absolute numerosity of one disk set while trying to disregard the other set, was adopted to investigate the task-specificity of the numerosity illusion. It is commonly believed that numerosity is represented in the form of analog magnitude that obeys Weber’s law (Moyer & Landauer, 1967; Whalen et al., 1999). In order to generate a linguistic label for the numerosity of a set of objects, a magnitude representation has to be built and then a
mapping process takes place to associate the numerosity magnitude with some appropriate symbols. Previous studies have established a compressive power function between the estimated numerosity and physical numerosity of a visual set (Krueger, 1972; Krueger, 1982). Since a numerosity estimation task is thought to tap the magnitude representation of numerosity, we hypothesize that, if the numerosity illusion discovered with a discrimination task indeed impinges on the numerosity representation, it should also be observable in an estimation task, i.e., subjects would give higher numerosity estimates for gray disks than for white disks. However, if the illusion instead reflects some kind of post-perceptual bias, it may disappear for absolute numerosity estimation.

A second purpose of this section is to examine whether the interaction between gray and white disks is sensitive to attention manipulation. In a discrimination task, subjects inevitably have to pay attention to both disk sets in order to compare their numerosity, which may trigger the interaction that is responsible for the numerosity illusion. In an estimation task, however, subjects only need to report the numerosity of one disk set. If they have knowledge about the target set beforehand, it is possible for them to attend selectively to that set while ignoring the other set, in which condition the interaction between the two sets of disks may diminish or disappear and accordingly the numerosity illusion is less likely to occur. To this end, in a random design, subjects were either post-cued or pre-cued as to the target set for estimation in an intermixed patch. A similar experiment was also run using a block design in an attempt to further ensure selective attention to only one set of disks. As a control, estimation performance was also measured with white disks and gray disks presented alone.
7.1 Experiment 10a: Post-cue Random Design – Intermixed

In this experiment, subjects estimated the numerosity of either gray or white disks on each trial and didn’t know which set to estimate until the disk display has disappeared. In order to perform the task, they had to pay attention to both disk sets and retain the numerosity of both. Therefore, this constitutes a condition that resembles a discrimination task in terms of attention allocation and in which, according to our hypothesis, the numerosity illusion is most likely to manifest.

7.1.1 Methods

The disk display was an intermixed patch of gray and white disks similar to that used in previous experiments. The relevant or target disk set, the one to be estimated on a trial, varied in numerosity in the range of 20-80 by steps of 10, while the irrelevant or distractor set had a fixed numerosity of 50. Therefore, there were two types of trials, one type on which the gray disks served as the target set and the white disks as distractor, the so-called ‘gray’ trials, and the other type on which the white disks served as the target set and gray disks as distractor, the so-called ‘white’ trials. These two types of trials were randomized throughout the experiment and there was no prior information for subjects to tell one from the other. Each trial proceeded as shown in Fig. 18. A fixation screen was first presented for 1s, followed by the disk display for 1.5s. After the disks disappeared, a verbal prompt which cued the subject to estimate the numerosity of either gray or white disks was provided and remained until subjects input their estimates using the number keys. The next trial began after a 3s inter-trial interval. Each numerosity level was repeated 10 times for each trial type throughout the experiment, which was divided into
four blocks, with 35 trials in each. Before the experiment, subjects completed 10 practice trials. Seven subjects participated in this experiment, which lasted about 40 minutes.

Figure 18. Flow of a numerosity estimation trial in Experiment 10. A trial begins with a fixation screen for 1000ms, followed by the disk display for 1500ms, after which a verbal prompt is given as to which disk set to estimate.

7.1.2 Results

The average performance of all subjects was plotted in Fig. 19A. Mean numerosity estimates of each condition were taken for individual subjects and subjected to a 7*2 ANOVA, with physical numerosity level and target color (white or gray) as repeated-measure variables. Not surprisingly, the main effect of numerosity was significant, F(6,36)=64.21, p<.01, suggesting that subjects were able to tell different numerosities apart and make estimation accordingly. Subjects’ numerosity estimate increased with the physical numerosity of the target set. The main effect of target color was also significant, F(1,6)=63.31, p<.01. As shown in Fig. 19A, subjects generally gave higher numerosity estimates for gray disks than for white ones, which was qualitatively consistent with the numerosity illusion reported earlier. The interaction between numerosity level and target
color was not significant, F(6,36)=.65, p=.69, suggesting that the trend to overestimate gray disks relative to white ones existed for the whole range of numerosities tested. It is also evident in Fig. 19A that the estimated numerosity was much compressed and can be described as a power function of the physical numerosity, consistent with previous findings (Krueger, 1972).
7.2 Experiment 10b: Pre-cue Random Design – Intermixed

Unlike in Experiment 10a, in this experiment, subjects were cued as to which disk set to estimate before the onset of the disk display, so they are motivated to attend only to the
target set while ignoring the distractor set. If the interaction between the two sets of disks is sensitive to whether attention is applied to both, due to pre-cuing, the chance of observing the numerosity illusion is much lowered in this experiment. However, to the extent that the interaction operates in an automatic way, the illusion should still occur.

**7.2.1 Methods**

The stimuli and procedure were identical to that of Experiment 10a, except that subjects were pre-cued regarding which target set to estimate on each trial. This was fulfilled by presenting either a gray or a white fixation cross at the beginning of the trial. Nevertheless, a verbal prompt congruent with the fixation cue was still provided after the disk display had terminated. Eight subjects participated in this experiment, which lasted about 40 minutes.

**7.2.2 Results**

One subject was an obvious outlier and was excluded from data analysis. The average performance of the remaining subjects was plotted in Fig. 19B. The pattern of results was quite similar to that of Experiment 10a. Again, a 7*2 ANOVA was performed on mean numerosity estimates, with numerosity level and target color as two repeated-measure variables. There was a significant main effect of numerosity, $F(6,36)=45.88, p<.01$, showing that subjects’ numerosity estimates increased with physical numerosity. The main effect of target color was again significant, $F(1,6)=30.88, p<.01$, with the numerosity of gray disks overestimated relative to white disks. The interaction between physical numerosity and target color was not significant, $F(6,36)=.71, p=.64$, suggesting a
general overestimation of the gray disks over the whole numerosity range. It is also worth noting that the outlier subject showed the same trend of overestimating the numerosity of gray disks, however, to a much-exaggerated degree.

7.3 Experiment 10c: Block Design – Intermixed

Attempts were made in Experiment 10b to manipulate the allocation of attention by subjects so that they attended only to the target disk set. The effectiveness of the pre-cuing procedure used was confirmed by the improvement in overall accuracy for the pre-cued estimates in Experiment 10b over the post-cued estimates in Experiment 10a. However, even though attention was directed at a single disk set by pre-cuing, the numerosity illusion was still evident: subjects consistently gave higher numerosity estimates for gray disks than for white ones at all numerosity levels. It is possible that in a random design, since subjects had to constantly switch between estimating the numerosity of gray disks and that of the white ones, they still had a strong incentive to pay attention to both and therefore the manipulation was not sufficient to suppress the interaction between gray and white disks. To further ensure that subjects selectively attend to one disk set and ignore the other, in this experiment a block design was adopted, in which subjects estimated the numerosity of the same disk set, either white or gray, throughout each block.

7.3.1 Methods

The stimuli and procedure were similar to that of Experiment 10a. However, a different range of numerosities, 30 through 70 in steps of 5, was used, thus requiring subjects to
make finer estimation than in both experiments above. Moreover, in this experiment ‘gray’ trials and ‘white’ trials were grouped into separate blocks. For each trial type, there were two consecutive blocks, each with 45 trials. Before the onset of each block, subjects were informed of the target set for the whole block. Accordingly, the verbal prompt delivered at the end of each trial remained unchanged throughout each block. Eight subjects participated in the experiment. Half of them completed the two ‘gray’ blocks first and then the ‘white’ blocks while the other half did the opposite order. The experiment lasted about 40 minutes.

7.3.2 Results
The average performance of all subjects was plotted in Fig. 19C. A 9*2 ANOVA was performed with numerosity level and target color as repeated-measure variables. The main effect of numerosity was significant, F(8,56)=50.18, p<.01. Despite the finer spacing of numerosity levels in this experiment, numerosity estimate by the subjects increased steadily with the physical numerosity. The main effect of target color was also significant, F(1,7)=6.80, p=.04, showing that the numerosity of gray disks was overestimated relative to the white ones. There was no interaction between target color and physical numerosity, F(8,56)=.64, p=.74. Overall, the pattern of results is highly similar to both Experiment 10a and 10b, suggesting that how subjects allocate attention is not a critical factor in the occurrence of the numerosity illusion.

7.4 Experiment 10d: Unmixed Condition
In the above experiments, a numerosity illusion similar to the one we reported earlier was consistently measured with a numerosity estimation task, regardless of the status of attention allocation. To make sure that we are measuring the same illusion as that with a discrimination task, in this experiment white disks and gray disks were presented alone without intermixing on each trial to simulate the segregation conditions investigated in previous sections. As demonstrated earlier, one prominent feature of the numerosity illusion is its vulnerability to segregation of any type between the two sets of disks. If indeed the estimation task and the discrimination task used in this study tap into the same underlying numerosity mechanism, no illusion should appear in the current experiment.

7.4.1 Methods

The disk display comprised a single set of either gray or white disks on each trial, with the numerosity varying in the range of 20-80 by steps of 10. Subjects were asked to estimate the numerosity of whichever disk set was presented. All other aspects were identical to that of Experiment 10a. Six subjects participated in this experiment, which lasted about 40 minutes.

7.4.2 Results

The average performance of all subjects was plotted in Fig. 19D. A 7*2 ANOVA was performed on the mean numerosity estimate with numerosity level and target color as repeated-measure variables. The main effect of numerosity was again significant, \( F(6,30)=68.59, \ p<.01 \), so that numerosity estimate increased with physical numerosity. However, the main effect of target color was not significant, \( F(1,5)=3.73, \ p=.11 \), showing
no difference in numerosity estimate between gray and white disks. The interaction between target color and physical numerosity was not significant, either, F(6,30)=1.18, p=.34. Therefore, when the two sets of disks were presented separately, in contrast to the experiments above, the numerosity of gray disks was not estimated as greater than that of white ones.

### 7.5 Discussion

In this section, attempts have been made to investigate the task-specificity of the numerosity illusion reported in earlier sections. It is found that, in an estimation task in which subjects had to verbally report the absolute numerosity of a set of disks, there existed a consistent tendency for subjects to overestimate the numerosity of gray disks relative to that of white ones in an intermixed patch of disks, which is qualitatively similar to the numerosity illusion we measured in a discrimination task. Critically, when the two sets of disks were presented separately, the relative advantage of gray disks in numerosity estimate disappeared, which is highly consistent with previous demonstrations that segregation in general destroys the numerosity illusion we discovered, thus suggesting that segregation similarly affects numerosity estimation and discrimination when it comes to perceiving the numerosity of an intermixed patch of disks as used in our experiments. All these results indicate that the interaction between the two disk sets that results in the relative overestimation of gray disks takes place at a relatively early stage tapped by both numerosity estimation and discrimination. Since the effect manifests itself in an estimation task where no explicit comparison between disk patterns is involved, it cannot be simply attributed to any post-perceptual comparative
biases. Rather, we suggest that the magnitude representation underlying numerosity perception is altered by the intermixing of gray and white disks.

Another piece of observation that supports the view that numerosity estimation and discrimination draw on the same underlying representation lies in the comparison between the results of the pre-cued experiment and the unmixed experiment (Fig. 13B and 13D). For both experiments, subjects had to attend to a single disk set, the difference being that the target set was presented in an intermixture for the former and in isolation for the latter, so they constitute a perfect pair for evaluating the effect of the presence of one disk set on the perceived numerosity of the other set. Although statistically insignificant, it is apparent from the comparison that the estimated numerosity of the gray disks remained unaffected while that of the white disks decreased when placed in an intermixture relative to when they were presented alone, which is again fully consistent with the results from a discrimination task, as demonstrated in Experiment 8.

By comparing results from different experimental designs, evidence was also gained that the interaction between the two disk sets in an intermixed patch seems to be mandatory. In an attempt to discourage subjects from processing the irrelevant disk set, we either pre-cued them about the target disk set in a random design or employed a block design where the target set remained unchanged throughout. Neither manipulation had any effect on the relative advantage of gray disks in terms of perceived numerosity, although there were signs that subjects showed an overall benefit in numerosity estimation under such manipulations. Therefore, the effect persists no matter how subjects allocate their
attention between the two disk sets, which points to a pre-attentive stage of processing as the locus of the effect. This is reminiscent of the notion put forward earlier that a mandatory segregation process takes place when trying to extract the numerosity of a subset of elements from an intermixture. It is during the segregation process, we suggest, that the interaction between different sets of elements arises, which eventually leads to the numerosity illusion observed in this study.
Chapter 8. General Discussion

8.1 Mechanism of the Illusion and Numerosity Computation

We have shown in the current study that a reliable numerosity illusion occurs when equal numbers of gray disks and white disks were intermixed on a dark gray field. The illusion can also be generalized to negative contrast. In both cases, the disk set with a lower absolute contrast appears to outnumber the set with a higher contrast. However, we were able to eliminate the illusion by segregating the two disk sets in various ways. They were either physically segregated into separate visual fields or perceptually segregated by stereopsis, by motion or by orientation. In all cases, a contrast difference per se failed to induce the illusion. We have therefore demonstrated the important role of segregation in general and the lack thereof in the occurrence of the illusion. We conclude that, whereas contrast itself provides a cue for segregation, it is not a sufficient one and at best leads to incomplete segregation, which provides the basis for the numerosity illusion to occur.

To account for the mere observation that there appear to be more gray disks than white disks in an intermixed patch even though they match in physical numerosity, we have considered several alternative explanations, ordered from low- to high-level mechanisms involved:

1. Asymmetrical sensory interaction may exist between gray/low-contrast and white/high-contrast disks. High-contrast elements, somehow, are suppressed more (or possibly boosted less) by the presence of low-contrast elements than the other way around towards the computation of numerosity.
2. Contrast is an effective depth cue (O'Shea, Blackburn, & Ono, 1994) and the resulting depth separation, white disks seen in front and gray disks at the back, induces the illusion. Why might depth matter then? For one thing, assuming that numerosity is inferred from continuous variables such as aggregate area, as a result of size constancy, gray disks seen at the back are perceived to be bigger than white disks in front and resulted in larger aggregate area hence more numerous. For another, real world objects tend to occlude each other. Even though no actual occlusion occurs in our stimulus, the visual system may infer hypothetical gray disks occluded by white disks, which are physically invisible yet estimated. We call this the occlusion hypothesis.

3. Confusion or misclassification occurs when disks that only differ in contrast are intermingled. Some white disks are confused or misclassified as gray ones but gray disks are less likely to ‘become’ white ones, resulting in a relative advantage for the gray disks.

4. A comparative bias exists that compensates for the effort of ‘counting’ gray disks, which are presumably more resource-consuming than that of white disks due to their lower visibility.

A series of experiments have been conducted in an attempt to prove or disprove each of these hypotheses as they make vastly different predictions. According to our data, the illusion cannot be a depth effect. In one of our experiments, the depth order of gray and white disks were manipulated using stereopsis so that in one condition the white disks were placed in front of the gray disks, thus reinforcing the contrast-induced depth
layering, and in another condition the gray disks were placed in front of the white disks, thus counteracting the depth impression induced by contrast. The depth hypothesis would predict the same or an even stronger numerosity illusion for the ‘white-in-front’ condition and predict a reversal in the illusion for the ‘gray-in-front’ condition, which was disproved by our data. It was found that depth separation, regardless of the depth order, simply abolished the illusion, consistent with the general role of segregation. Further disproof came from the matching experiments. It was shown that the perceived numerosity of white disks actually decreased when presented in an intermixture compared to when presented alone whereas that of gray disks remained unchanged. This is in direct contradiction to a depth effect, in particular to the occlusion hypothesis, which would instead predict an increase in the perceived numerosity of gray disks and a preservation of that of white ones. While under certain conditions depth may indeed have an effect on numerosity judgment, as documented in previous studies (Schutz, 2012; Tsirlin et al., 2012), the numerosity illusion we reported here can’t be ascribed to a depth effect.

We have also disproved the misclassification hypothesis. If misclassification is all that happens among the disks, an obvious prediction is that the total number of disks in the patch should be preserved and a reduction in the perceived numerosity of white disks should result in a corresponding increase in that of the gray ones. Our data, however, didn’t agree with this prediction. We showed in a superset matching experiment that there was a loss in the total numerosity of white and gray disks, consistent with the results of matching each set separately. Apparently, some white disks were lost without converting
into gray disks. Stronger evidence against the misclassification hypothesis came from matching with noise. When one set of disks that served as the source of influence was simulated with equivalent contrast noise, owing to the drastic difference in appearance of the target disks and the distractors, misclassification was not any more possible yet the effect remained. Therefore, misclassification does not constitute a valid explanation for the numerosity illusion.

Can the illusion arise from a biased comparison process? If yes, we would predict that when no comparison is required between the two disk sets, no illusion or similar effect should occur. However, this turned out not to be true. In the matching task we used, subjects only needed to discriminate the numerosity of disks of the same color (contrast) between two patches, thus no explicit comparison between the two colors was involved. Nevertheless, an effect was obtained that was perfectly consistent with the illusion both in direction and in magnitude, i.e., the white disks decreased in perceived numerosity by an amount of 5-10 disks while the perceived numerosity of gray disks was largely preserved. We have further tested the task-specificity of the effect by engaging subjects in a numerosity estimation task in which the subjects had to estimate the absolute numerosity of only one disk set on each trial. In some conditions subjects were always informed of the target set beforehand so as to discourage them from attending to the other set, thus making the distractor set totally irrelevant to the task. However, a similar effect indicating the illusion was observed for all conditions as long as the two disk sets were intermixed. Therefore, the numerosity illusion doesn’t seem to be specific to the
discrimination task and can’t be simply attributed to a comparative bias at the response stage.

So far, our data favors the existence of an asymmetrical sensory/perceptual interaction between elements of different contrasts that occurs at a relatively early stage of processing. Whereas no attempt is made to figure out the exact mechanism underlying the computation of numerosity, we here tentatively propose a conceptual model to help explain our findings. We call it the thresholding hypothesis (Fig. 20). Concurring with Dakin et al. (2011), we posit that the computation of numerosity is not based on a token system that tags each individuated item on a location map but instead relies on the spatial frequency content of an unsegmented raw image. Numerosity is indexed by the contrast energy conveyed in the spatial frequency channel appropriate to the scale of the stimulus. Ideally, a numerosity mechanism would be invariant to the contrast of the stimulus. To derive contrast invariance, the overall contrast energy is normalized by the prevailing local contrast. This framework works fine with most ordinary situations, e.g., when applying to a homogenous set of objects. However, for an intermixed patch of white and gray disks, to compute the numerosity of each set, a thresholding process is required to split them apart. As shown in Fig. 20A, the model consists of three stages. The input image is first subjected to conventional spatial filtering, after which each disk is represented by a Gaussian profile, with white disks higher in amplitude than gray disks. The local maximum of this neural image thus follows a bimodal distribution, two modes corresponding to the peak responses to gray and white disks, respectively. The thresholding process then takes place, producing two separate images as the basis for
Figure 20. The thresholding model. (A) Stages of numerosity computation: spatial filtering, thresholding and normalization. See text for details. (B) After thresholding, the effective contrast energy of gray disks (blue slashed region) remains unaffected while that of white disks (orange slashed region) is reduced. Each Gaussian curve represents the neural responses to a single disk.

Contrast energy measurement. To derive the image of gray disks, neural responses above a properly chosen threshold are suppressed; the threshold is ideally the peak response to gray disks. Therefore, the white disks are effectively removed from the image and do not interfere with the extraction of contrast energy and subsequent normalization to derive
numerosity metrics. Although we do not specify the exact mechanism by which the white disks may be suppressed, it resembles the process of perceptual scission (Anderson & Winawer, 2008). To derive the image for white disks, however, the gray disks are treated as part of the background and the neural noise contributed by them needs to be filtered out. This is done by thresholding the original image at the same threshold as above, but now resulting in an elevated background level. As a result, the effective contrast of the white disks is reduced (Fig. 20B). The contrast energy gauged from the thresholded image is then normalized by local contrast present in the original image to derive the numerosity metric which, suffering from a loss in effective contrast energy, is responsible for the relative underestimation of the numerosity of the white disks.

This model is therefore able to explain our major findings, in particular the fact that the perceived numerosity of white disks was decreased while that of gray disks was preserved in an intermixed patch. It also explains why contrast noise could have equivalent effect to that of a distinct set of distractors on the perceived numerosity of the target disk set. This is simply because the same thresholding process operates whether the target set is intermixed with another set of distractors or is embedded in visual noise, as long as the latter two elicit similar contrast response. To account for the results of the segregation experiments, we just need to assume that the thresholding process takes place after some level of scene analysis, which in our view is a very plausible assumption. Once the two disk sets are completely segregated, thresholding is no more necessary, hence the absence of the illusion. Several testable predictions can be borne out of this model. The first one regards the amount of contrast of the gray disks. The model predicts that as the contrast of the gray disks increases (but still lower than that of the white disks),
the threshold should be raised, which would result in more severe loss in effective contrast of the white disks, thus strengthening the numerosity illusion. However, the results of our Experiment 3 seemed to be inconsistent with this prediction. In that experiment, two different gray levels were used but no significant difference in illusion magnitude was observed. Since that experiment is not designed specifically to examine the effect of contrast levels, it is not clear if the violation is due to lack in statistical power or some other constraints. To the contrary, a pilot experiment using several more contrast levels has produced data in support of the prediction. Apparently, a more systematic investigation is required to resolve this conflict. Another prediction concerns the contrast polarity of the two disk sets. The model predicts that if the two disk sets have opposite contrast polarity, one with a luminance value above and the other below the background level, the numerosity illusion should disappear, even though there is still a difference in contrast magnitude. In such a situation, the intermixed image would be thresholded most effectively at zero contrast, i.e., a half-wave rectification is sufficient to split the image, which would have no effect on the calculation of contrast energy of either target set. This situation can also be viewed as another instance of complete segregation, since two disk sets with opposite contrast polarity are presumably easy to segregate (Grieco, Casco, & Roncato, 2006). To gain further evidence to bolster the model, future studies should put these predictions to test.

There has been much controversy as to how numerosity is perceived. Some researchers advocate that numerosity is a primary visual sense and is perceived directly and independently of other known visual dimensions (Burr & Ross, 2008; Dehaene, 2011).
Others argue that numerosity is at best a secondary visual dimension and should be inferred from measures of other primary dimensions such as density (Dakin et al., 2011; Durgin, 2008; Gebuis & Reynvoet, 2012a). In rough correspondence, when constructing numerosity models, some have introduced a token system or a location map and used individuated objects as the input (Dehaene & Changeux, 1993; Verguts & Fias, 2004), whereas others have disputed the concept of tokens and instead used unsegmented raw image as input (Dakin et al., 2011; Morgan et al., 2014). It is worth noting that the occupancy model proposed by Allik and Tuulmets (1991), while seeking surrogate measures (which is occupied area in their case) for numerosity, has implicitly relied on a token system by assuming that each individual object has its own territory. None of these models in their current form can account for our data. These models with a reliance on object tokens, although having received empirical support (Franconeri et al., 2009; He et al., 2009), are particularly problematic in this regard. Our occlusion hypothesis and misclassification hypothesis have both implicitly assumed the existence of object tokens, but both were immediately rejected by the data. The occupancy model has the potential to explain our results, with the additional assumption that occupancy radius of white disks somehow shrinks in the presence of gray disks, which, however, is itself in need of an explanation and may have to resort to some low-level mechanism similar to the kind of sensory/perceptual interaction we have proposed. To the contrary, we are convinced that the other type of models that operate on raw image have better chance of capturing the major features of our results. Therefore, our conceptual model resembles that proposed by Dakin et al. (2011) and Morgan et al. (2014). Whereas we do not rely on the specific mechanism they devised for the calculation of numerosity, our model has borrowed from
them the concept of contrast energy as the basis for deriving metrics for numerosity. Unlike theirs, which treat the whole set of objects as the enumeration target, our model has provided a mechanism for extracting the numerosity of a subset of objects defined by luminance contrast from an intermixture. We have further demonstrated that segregation of a visual image at the set level rather than at the object level is important for numerosity computation, the failure of which may lead to non-veridical perception, as exemplified by our illusion. Although we favor the raw image as the immediate input to numerosity computation, we admit that there may still be a role for a token system to take. It is possible that at some point after the basic metrics for numerosity have been derived, the token system comes into play and accounts for the conscious impression that in a numerosity task we are actually estimating the numerosity rather than some other image statistics.

8.2 Saliency/Attention, Perceptual Grouping and Numerosity

One most counterintuitive aspect of the numerosity illusion is that weaker stimuli that have lower contrast win over stronger stimuli that have higher contrast in perceived numerosity. The counterintuitive nature of this illusion is further exacerbated by the fact that the perceived numerosity of the low-contrast stimuli appears to be preserved while that of the high-contrast stimuli appears to be reduced by intermixing, compared to when they are presented alone. These results are counterintuitive for at least two reasons. Firstly, low-contrast stimuli are somewhat less visible than high-contrast ones thus more difficult to individuate from the background for enumeration. To the extent that individuation is a necessary step in the computation of numerosity, while both sets may
be subject to underestimation relative to their physical numerosity (Krueger, 1972), we would predict that weaker stimuli should suffer more. Secondly, stronger stimuli are also more salient thus attracts attention, which would potentially facilitate the ‘counting’ process and may result in a benefit in perceived numerosity relative to weaker thus less salient stimuli (Trick & Pylyshyn, 1994). Our findings have important implications for the role of salience as well as that of attention in numerosity perception.

Tokita and Ishiguchi (2010a) investigated the effect of salience on perceive numerosity. In their study, different pairs of elements were intermixed in the way similar to our stimulus. The pair could be red and green dots, which constitute a symmetrical pair with regard to search efficiency in a visual search task. Other pairs, however, are search asymmetrical, consisting of circles and circles with lines or of circles and circles with gaps. For the latter case, circles with lines and circles with gaps are said to pop out in visual search and are deemed more salient than perfect circles. It was found that subjects overestimated the relative numerosity of pop-out elements in an intermixed patch. By contrast, no overestimation in either direction was observed for the symmetrical pair of red and green dots. Therefore, it seems that salient stimuli tend to appear more numerous, presumably due to the attentional priority they receive over less salient stimuli.

In sharp contrast, our findings have revealed an opposite trend. In our study, gray and white disks were used that also constitute a search asymmetrical pair in which the white disk should pop out (Treisman & Gormican, 1988). In our case, however, a disadvantage was observed instead for the more salient white disks with regard to perceived
numerosity. Apparently, stimulus saliency can have opposing effects on numerosity perception, depending on the specific stimuli used. One tentative explanation that can possibly reconcile the conflicting findings is perceptual grouping. We hypothesize that perceptual grouping reduces perceived numerosity, perhaps mediated by the increase in perceived clustering (Ginsburg & Goldstein, 1987). In an intermixed patch of gray and white disks, the stronger stimuli, which are the white disks, tend to group more than the weaker gray disks, resulting in a reduction in relative numerosity. In the stimuli used by Tokita and Ishiguchi (2010a), it is plausible that the circles, which are the simpler of the pair, group to a larger degree than the more complex circles with lines or circles with gaps, which may potentially account for the relative numerosity benefit for the latter. Besides, stimulus complexity itself may contribute to numerosity perception, producing the effect they observed. Therefore, stimulus salience might just be a confounding factor in their study as well as in ours, since it is presumably the stimulus strength rather than salience that really matters. In any event, stimulus salience or attention thus attracted may not be a critical factor in generating our numerosity illusion. This notion is consistent with Burr et al. (2010), who demonstrated that the approximate number system is insensitive to attentional load, and is further corroborated by the findings from our numerosity estimation experiments, where the manipulation of attention hardly had any effect on the illusion.

Whereas perceptual grouping provides a plausible explanation for our illusion, questions remain as to why grouping between white disks should be enhanced relative to that between gray disks when they are intermixed and how this interaction feeds into the
computation of numerosity so as to produce the illusion. It has been found in many studies that clustering reduces perceived numerosity (Ginsburg, 1991; Ginsburg & Goldstein, 1987; Koesling et al., 2004), so one possible avenue for perceptual grouping to impact perceived numerosity is by increasing the perceived degree of clustering. However, how clustering in turn affects perceived numerosity is itself in need of a mechanism (Vos et al., 1988). On the other hand, as Durgin (1995) has pointed out, perceived clustering might be as well just one consequence of numerosity perception rather than the source of it. To avoid this circularity, a more simplistic account, the perceptual suppression or thresholding account, has been advocated instead to explain the numerosity illusion, which is able to capture the main thrust of our data and is preferred over the perceptual grouping account.

8.3 Contextual Effect in Numerosity Perception

The numerosity illusion reported in this study can certainly be construed as an instance of contextual effect. We showed that the perceived numerosity of one element set is affected by the presence of another set when the two sets are intermixed. Cordes et al. (2014) has demonstrated a similar contextual effect. In their study, two sets of differently colored disks were intermixed and subjects had to estimate the numerosity of one disk set while ignoring the other set. It was found that numerosity estimation of the target set was decreased in the presence of a distractor set compared to when the target set was presented alone without distractors. Our findings partly resemble theirs in that the perceived numerosity of white disks appears to have decreased when they are intermixed with gray disks. However, the important departure from theirs is the asymmetry of the
effect we have observed: the perceived numerosity of gray disks remains unaffected by the presence of white disks. We speculate that the discrepancy is due to the nature of the different stimuli used in the two studies.

Cordes et al. (2014) has used disks of different colors, which makes it easy for the two disk sets to segregate from each other. Once segregated, numerosity can be extracted independently for each set, which we have also observed for our stimuli by using strong segregation cues. The contextual effect they reported may have occurred at a later stage and reflect a normalization process at the level of numerosity representation. Consistent with this idea, in their study, the amount of numerosity underestimation was found to be an inverse function of the size of the distractor set. By contrast, our stimuli were composed of disks of different contrast levels. Conceivably, segregation based on contrast is incomplete (Beck et al., 1991), and as a result, enumeration of the two disk sets cannot be carried out independently of each other, allowing for the asymmetrical effect to occur by the sort of mechanism we have described in other section. We have further demonstrated that the influence of one disk set on the other set can be simulated using equivalent visual noise, without the explicit existence of a comparison disk set as the source of influence. Clearly, a numerosity normalization process is not any more plausible with such a stimulus configuration yet the effect remains. Therefore, our numerosity illusion should best be construed as a contextual effect occurring at a relatively early stage of processing that contributes to later numerosity computation. This also explains why the same illusion was observed with both relative numerosity discrimination and absolute numerosity estimation.
Other types of contextual effect have also been reported in literature. For example, Bevan et al. (1963) had subjects estimate the numerosity of beans in a jar. It was found that estimation depended on the size of the jar and on whether it was viewed as part of the figure with the beans or as the background. However, in their study, they used only verbal instruction to manipulate the figure-ground assignment, so the contextual effect they observed may have reflected a response bias rather than an effect on perceived numerosity itself. Contextual effects have been common in visual processing (Albright & Stoner, 2002), and it is not surprising at all that it also exists in the domain of numerosity perception. However, the specific mechanisms underlying a contextual effect may differ, depending on the stimulus and the task used.

8.4 The Role of Contrast In Visual Processing

Contrast is the difference in luminance between adjacent regions of a visual image. The visual system is by design more sensitive to contrast than to absolute luminance. As early along the visual pathway as the retina, receptive fields of many neurons comprise antagonistic regions that respond respectively to light and dark stimulation, providing an ideal structure for analyzing contrast. For example, both retinal ganglion cells and lateral geniculate nucleus (LGN) neurons have a center-surround structure, so they respond best to a light or dark spot that fills the center of their receptive fields. Simple cells in the striate cortex have elongated receptive fields, with an excitatory central oval subregion, for example, flanked by inhibitory subregions. This type of arrangement renders the cells
sensitive to bars of a particular orientation, which is a desirable property for edge detectors.

Since visual information in an image is largely conveyed by contrast, how it affects visual processing is an important question to ask. It is found that contrast sensitivity depends on spatial frequency, with the peak situated around 4-6 cycles/degree in the intermediate range (Campbell & Robson, 1968). Spatial frequency channels with differential contrast sensitivity have been posited to underlie this dependence on spatial frequency, which has received considerable empirical support, as demonstrated in selective adaptation and masking. According to this view, a visual image is subject to spatial filtering at multiple scales before any further processing takes place. Since these spatial filters necessarily respond to contrast information in an image, any further stages of visual analysis that utilize the output of these filters may potentially be affected by contrast, depending on the specific mechanisms involved.

At the neuronal level, response of visual neurons typically increases with increasing contrast before finally saturating at high contrast (Albrecht & Hamilton, 1982). Contrast also affects the response latency of many visual neurons, which increases with decreasing contrast (Oram, 2010). Spatial summation by visual neurons also depends on contrast. Both in the primary visual cortex (V1) and in the middle temporal area (MT), neurons have been found that increase in summation area as stimulus contrast decreases, so that these neurons respond more strongly to a large stimulus of low contrast than to one of high contrast (Pack, Hunter, & Born, 2005; Sceniak, Ringach, Hawken, & Shapley, 1999).
This dependence on contrast of spatial summation has its perceptual consequence. For example, it has been demonstrated in human subjects that discrimination of motion direction is actually easier for a large moving dot pattern at low contrast than one at high contrast (Tadin, Lappin, Gilroy, & Blake, 2003). These results suggest that contrast is an important stimulus parameter that has a profound effect on visual processing.

While contrast is a fundamental property of any visual image, in many circumstances, it is the structural information (form and its motion) conveyed by contrast that is perceptually relevant. Ideally, some aspects of visual processing should be contrast-invariant, especially as the abstract level of a perceptual goal increases. An extreme example is object recognition, which, for it to be useful, should for the most part be invariant with changes in viewing conditions, including that in stimulus contrast. Indeed, it has been found that in higher-order visual areas that are responsible for face and object recognition, neural activity is largely insensitive to image contrast (Avidan et al., 2002). Some lower-level visual processing, such as orientation and spatial frequency selectivity, has also been found to be mostly independent of contrast for suprathreshold stimuli (Sclar & Freeman, 1982; Skottun, Bradley, Sclar, Ohzawa, & Freeman, 1987). To achieve contrast invariance, some type of normalization process seems to be necessary (Carandini & Heeger, 1994, 2012). However, this may fail or alternatively be misapplied in certain situations. For example, perceived motion speed seems to depend on contrast. A moving stimulus of lower contrast typically appears to be slower than one of higher contrast, but this may reverse under some conditions (Blakemore & Snowden, 1999; Stone & Thompson, 1992). Accordingly, the motion direction of a plaid pattern also depends on
the contrast of its component gratings, as if a contrast-based weighting of speed is applied before the direction of the pattern motion is determined (Stone, Watson, & Mulligan, 1990).

Although the role of contrast in visual processing has generally been considered within a feedforward approach, it may also operate by some high-level feedback mechanisms. For example, contrast has been suggested to be an effective depth cue, with lower-contrast stimuli appearing to be farther away from the observer than higher-contrast ones (O'Shea, Blackburn, & Ono, 1994; O'Shea, Govan, & Sekuler, 1997). Consistent with this potential role of contrast, it has been demonstrated that a moving plaid pattern composed of two component gratings differing in luminance contrast and moving in different directions tends not to cohere as pattern motion, but instead component motion is more often perceived, presumably due to the depth separation between the two component gratings implied by the contrast difference (Stoner & Albright, 1998). These findings thus suggest a role for contrast in surface segmentation, which in turn affects other stages of visual processing.

The basic result of this study that contrast affects numerosity perception but only when elements of different contrasts are unsegregated has implications for both the image-based and the scene-based approaches to visual processing discussed above. In an intermixed patch of disks with two different contrasts, to extract the numerosity of each disk set, one needs to first divide the whole population of disks into two lightness-defined subpopulations. While in a very different context, Beck et al. (1991) used a similar
arrangement of stimulus to that used in this study and proposed the concept of population segregation, which they found to be a function of the lightness (contrast) difference between the two element sets. Concurring with them, we hypothesize that the so-called population segregation based on lightness (or contrast) is far from complete, which provides the basis for the numerosity illusion to occur. While the specific mechanism underlying the illusion is yet unclear, our data can be best explained by a perceptual suppression account which we tried to formulate using an image-based approach. However, scene-based mechanisms clearly played a role in our experiments when segregation was facilitated by other cues that were stronger than contrast and blocked the illusion. We demonstrated that, once the segregation is complete, numerosity perception turns out to be largely contrast-invariant. Therefore, like the computation of other visual dimensions, perhaps with the best analogy to be found in visual motion analysis, the apprehension of numerosity seems to be achieved through a close synergy between feedforward and feedback mechanisms.
References


