Perceived magnitude of visual displays: Area, numerosity, and mean size

Hunjae Lee
Graduate Program in Cognitive Science, Yonsei University, Yonsei University, Seodaemun-gu, Seoul, Korea

Jongsoo Baek
Yonsei Institute of Convergence Technology, Yonsei University, Yonsei University, Seodaemun-gu, Seoul, Korea

Sang Chul Chong
Graduate Program in Cognitive Science, Yonsei University, Yonsei University, Seodaemun-gu, Seoul, Korea
Department of Psychology, Yonsei University, Yonsei University, Seodaemun-gu, Seoul, Korea

Previous studies have shown that the visual system is able to estimate properties such as area, numerosity, and mean size efficiently and accurately. In the current study, we investigated whether our percepts of each of them could be based on ratios of the other two of these three properties. In each trial, observers viewed a display containing various quantities of filled circles and judged whether the magnitude of a property of the display, such as summed area, numerosity, or average size of the circles, was greater or less than a corresponding probe display. We found that mean size judgments were more accurate and precise compared to the other judgments. We then predicted observers' performances for each task using the measured performance for the other judgments. The results showed that the other properties predicted perceived summed area, but not perceived mean size and numerosity. Together, our results suggest that the visual system does not use ratios to compute mean size and numerosity.

Introduction

Ensemble perception has attracted much research interest because it can provide an efficient summary of a complex scene (Alvarez, 2011; Ariely, 2001; Chong & Treisman, 2003). A typical scene contains highly structured and redundant information (Kersten, 1987). One way of utilizing this structure and redundancy is to represent various features and objects in a scene as a statistical summary, such as mean size. Thus, it is important to study how the visual system computes the statistical properties of a scene.

One step toward understanding the mechanism of computing the statistical properties of a scene is to investigate whether the visual system computes a specific statistical property based on other properties. For example, people could derive the mean size of visual arrays using other properties such as summed area and numerosity. If the visual system represents summed area and numerosity, it can simply take the ratio of summed area to numerosity to derive mean size. It would be inefficient for the visual system to represent all three properties if one of them can be computed based on the other two properties.

The visual system represents not only absolute quantity but also ratios of quantities (Jacob, Vallentin, & Nieder, 2012). Single neurons in the prefrontal and parietal cortices are tuned to preferred proportions (Vallentin & Nieder, 2008, 2010). Similar results were found in human frontoparietal regions using an fMRI adaptation paradigm (Jacob & Nieder, 2009b). Based on these neurophysiological findings, Jacob et al. (2012) suggested that the visual system is able to make analog representations of magnitude ratios. Since the visual system is capable of representing magnitude ratios, it is possible that the visual system computes mean size based on the ratio of summed area to numerosity.

Citation: Lee, H., Baek, J., & Chong, S. C. (2016). Perceived magnitude of visual displays: Area, numerosity, and mean size. *Journal of Vision, 16*(3):12, 1–11, doi:10.1167/16.3.12.

doi: 10.1167/16.3.12 Received December 10, 2014; published February 12, 2016 ISSN 1534-7362

However, there has not been an empirical investigation of whether the visual system uses this ratio.

On the other hand, direct computation of mean size of visual displays might be more efficient and advantageous than derivation of the ratio of summed area to numerosity when the internal representations of the summed area are highly correlated with that of numerosity. It is inefficient to represent two correlated properties separately. In a visual scene, summed area is often correlated with numerosity because the total size of objects dramatically increases as the number of objects increases. Moreover, recent studies have suggested that there are, indeed, interrelationships between the representation of numerosity and other quantities. For example, in a meta-analysis, Walsh (2003) proposed a "generalized magnitude system" that processes time, space, and numerosity based on behavioral evidence (Brown, 1997; Casini & Macar, 1997; Dehaene, Dehaene-Lambertz, & Cohen, 1998) and the following neuroimaging studies. Neuroimaging studies have shown that the same posterior parietal area is responsive to the three properties (time: Leon & Shadlen, 2003; Onoe, Komori, Onoe, Takechi, Tsukada, & Watanabe, 2001; space: Stein, 1989; numerosity: Sawamura, Shima, & Tanji, 2002). In addition, activated regions in the intraparietal sulcus were overlapped during comparison tasks involving size, luminance, and numerosity (Pinel, Piazza, Bihan, & Dehaene, 2004).

Given the high correlation between numerosity and size, it is inefficient to represent numerosity and size separately. The visual system may derive numerosity from the ratio of summed area to mean size. Likewise, it may derive summed area from multiplying mean size by numerosity. Whether the visual system has a separate mechanism of computing numerosity has been highly debated (Burr & Ross, 2008a; Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Durgin, 2008). Because observers adapted to the number of items in a visual display, Burr and Ross (2008a) suggested that the visual system senses number directly. However, Durgin (2008) showed that number adaptation was determined by density, rather than numerosity. Dakin et al. (2011) also suggested that both number and density perception are based on a common metric, which is the ratio of responses tuned to low and high spatial frequencies. Therefore, it is important to investigate that the computation of visual magnitudes is based on the ratio of other properties.

In the present study, we investigated whether people perceive visual properties using the interrelationships among three specific properties: summed area, numerosity, and mean size. We first measured observers' ability to estimate these three properties using three different tasks: observers estimated the sum of stimulus sizes, numerosity of the stimuli, and their mean size. In each trial, observers viewed standard displays containing various quantities of filled circles and judged whether the magnitude of the standard display was greater or less than that of a subsequent test display. We then constructed psychometric functions describing the observers' performances for each task. Using these psychometric functions, we computed the point of subjective equality (PSE) to compare bias of each judgment and just-noticeable difference (JND) to compare sensitivity of each judgment. If one judgment has less bias and higher sensitivity than the other judgments, this judgment is not likely to be derived from the other judgments.

In the next step, we tested whether our percepts of each of mean size, summed size, or numerosity could be based on ratios of the other two of these three properties using the following modeling approach. For the mean size task, we first considered the mean size task's psychometric function as a performance-limiting noise distribution by converting a measured psychometric function (cumulative Gaussian) into a Gaussian distribution. We made performance-limiting noise distributions for the summed area and numerosity tasks separately using the same method. Then, we randomly sampled one value from the performance-limiting noise distributions from the summed area task and one value from the performance-limiting noise distributions from the numerosity task. We then took the ratio of the two values and repeated this process 1,000,000 times to make bootstrapped distributions of mean size perception. Next, using the bootstrapped distributions, we predicted the mean size task's performance. Similarly, the ability to estimate the summed area was predicted by multiplying mean size by numerosity and the ability to estimate numerosity was predicted by taking the ratio of summed area to mean size. If the ability to estimate one property is predicted by the other properties, predicted values should be positively correlated with measured values. Ideally, the slope of a regression line between the measured and predicted values should be one.

To preview the results, we found that mean size discrimination is more precise than summed area and numerosity discrimination. The predicted results showed that summed area was explained well by the other two properties, but both mean size and numerosity were not successfully described by the other two properties. Because mean size perception was more precise than summed area and numerosity and was not predicted by the ratio of summed area to numerosity, we suggest that the visual system does not perceive mean size based on the ratio of summed area to numerosity. Moreover, the visual system does not perceive numerosity based on the ratio of summed area to mean size because numerosity was not predicted by the ratio of summed area to mean size.

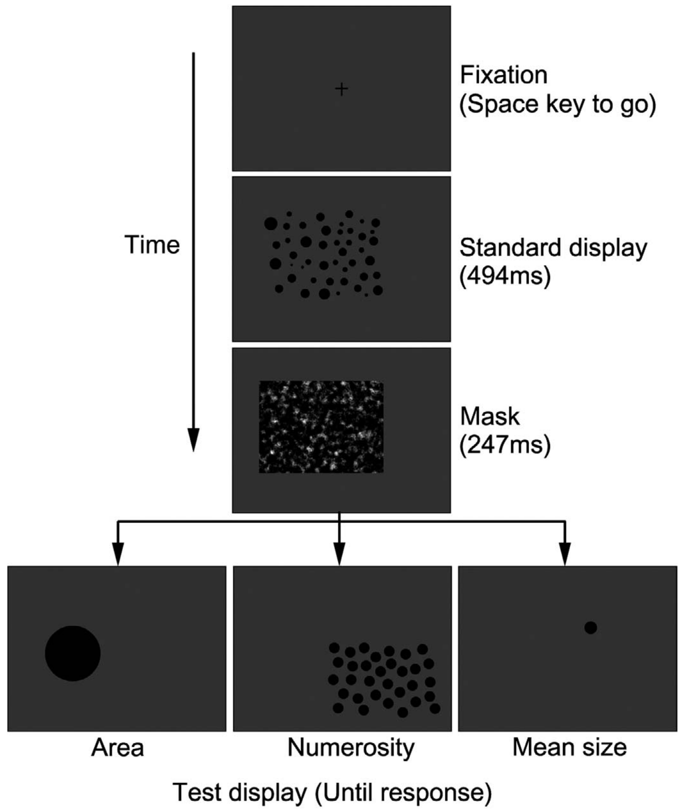

Figure 1. Stimuli and procedure for the experiment. For each task, observers were instructed to judge the perceived magnitude of a test display compared to a standard display.

Methods

Participants

There were four participants, including the first author, in this study. All had normal or corrected-to-normal visual acuity and were naïve to the purpose of the experiment, except the first author. All aspects of the study were carried out in accordance with the regulations of the Institutional Review Board of Yonsei University.

Apparatus and stimuli

Stimuli were presented on a 21-in. HP P1230 CRT monitor with a 1600 × 1200 pixel resolution using a refresh rate of 85 Hz. A chin and forehead rest was used to stabilize the head of each participant. The viewing distance was 90 cm, and thus, a pixel was 0.016°. Stimuli were generated using MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

As shown in Figure 1, there were standard, mask, and test displays. The standard display contained a set of black (0.01 cd/m^2) filled circles on a gray (14.43 cd/m^2) background. The diameter of each circle, ranging from 0.67° to 1.68°, was randomly selected. As Teghtsoonian (1965) found that the perceived size of a circle was related to its area by a power function with an exponent of 0.76, we generated circle sizes that were spaced equally on this scale. We converted each circle size and calculated the summed area and mean size on the scale (Chong & Treisman, 2003; Teghtsoonian, 1965). We randomly selected a set of circles repeatedly until the mean size of the set matched the mean predetermined for each condition. Because this procedure required time to generate each display, we pregenerated the sizes of each circle for all trials within a session. We used a mask, a phase-scrambled version of each display, to control the duration of a display (Ganis & Kutas, 2003). The test display contained the predetermined number of circles for numerosity judgments, but only one circle for both area and mean size judgments.

Design

There were four within-participant variables. The first was the type of task. We had three tasks: judging the sum of circle sizes, the numerosity of circles, and the mean size of circles. The remaining variables—density, numerosity, and mean size—were used to generate different displays. Each variable had three levels. Specifically, the sizes of the display fields (density) were 19.02° × 13.22°, 22.06° × 16.56°, and 25.6° × 19.2°. The width and height of the display field differed by 16% between the adjacent levels. For numerosity, the numbers of circles were set at 10, 20, and 40 to create enough room for generating test stimuli. The mean sizes were 1.02°, 1.15°, and 1.31° and the difference between the adjacent levels was 16% on the psychological scale (Teghtsoonian, 1965). In sum, participants performed three tasks, each of which had three different manipulations (area, numerosity, and mean size), each varying in three levels. This design generated 3 tasks × 3 levels of density × 3 levels of numerosity × 3 mean sizes, for a total of 81 conditions.

Participants judged whether the perceived magnitude observed for the test display was greater or less than that for the standard display. The standard display for each task had 27 different levels (3 different densities × 3 levels of numerosity × 3 different mean sizes). The perceived magnitude compared in the test displays differed across five different levels: 8% smaller and larger than the standard, 16% smaller and larger than the standard, and equal to the standard. Note that we rounded numbers in the case of numerosity variation. The total number of trials was 24,300 (81 conditions × 5 test displays × 60 repetitions), which were divided into 54 blocks of 450 trials each. Both the trial sequence in a block and the sequence of blocks were randomly determined. Before starting each session, participants performed 30 practice trials.

Procedure

Each trial was self-paced and started with a fixation point. When participants pressed the space bar, the standard display was presented for 494 ms. Immediately after the presentation of the standard display, the mask was presented for 247 ms, followed by the test display until a response was recorded. Participants judged whether, compared to the standard display, the perceived magnitude of the test stimuli was larger or smaller in the summed area and mean size tasks, and whether it was more or less numerous in the numerosity task. When they thought that the test display had a smaller magnitude, they pressed the 1 key. Otherwise, they pressed the 2 key. After the response, they moved on to the next trial by pressing the space bar. A 1-min break was given after 150 trials. Feedback was not provided, except during practice, to prevent bias (Bauer, 2009).

Results

Estimation biases: Point of subjective equality analyses

Psychometric functions were obtained for each condition and each observer by fitting cumulative Gaussian functions to the data. The functions of a representative observer are shown in Figure 2. To investigate the effect of three independent variables (density, number, and mean size of the displays) on each task separately, we collapsed the other two variables in Figure 2, which presents the psychometric functions dependent on each independent variable. We first analyzed the average point of subjective equality (PSE) for each task to determine the ability to estimate each magnitude. One-sample t tests showed that the PSE of the summed area (−3.98%) and mean size (1.46%) tasks did not differ significantly from zero, $t(3) = -1.687$, $p = 0.190$; $t(3) = 1.982$, and $p = 0.142$, respectively, suggesting that neither estimation was biased. However, the PSE of the numerosity task (−5.5%) was significantly less than zero, $t(3) = -3.931$, $p = 0.029$, suggesting that the observers underestimated numerosity. This underestimation may have been due to size variation in the standard display and uncorrelated sizes in the test display; such variation did not exist in previous studies (Allik & Tuulmets, 1991; Burgess & Barlow, 1983; Ross, 2003). Numerosity estimations in previous studies were conducted with uniform dot size in both the standard and test displays.

Figure 3 shows the PSEs of the three tasks depending on the three display types. First, the effect of density did not significantly influence any of the three tasks, $F(2, 6) = 0.386$, $p = 0.696$ for summed area; $F(2, 6) = 0.397$, $p = 0.689$ for numerosity; $F(2, 6) = 2.140$, $p = 0.199$ for mean size; these results are consistent with previous studies (Chong & Treisman, 2005; Burr & Ross, 2008a, 2008b). Second, regarding the effect of numerosity, the PSEs of the area task (the dotted line with squares) decreased significantly as the numerosity of the display increased, $F(2, 6) = 14.510$, $p = 0.005$, indicating that the summed area was underestimated as the numerosity of the display increased. Likewise, the PSE of the numerosity task (the line with triangles) decreased as the numerosity of the display increased, although the decrease was not significant, $F(2, 6) = 4.190$, $p = 0.073$).[1] In contrast, the PSE of the mean size task (the dotted line with filled circles) significantly increased as the numerosity of the display increased, $F(2, 6) = 15.531$, $p = 0.004$. Observers overestimated mean sizes as the numerosity of the display increased. More interestingly, while mean sizes were overestimated, perceived magnitudes of both summed area and numerosity were underestimated. A similar pattern was observed in the mean size variation. The PSEs of the summed area and mean size task significantly decreased as the mean size increased, $F(2, 6) = 16.473$, $p = 0.004$ for summed area; $F(2, 6) = 19.090$, $p = 0.003$ for mean size, indicating that the summed area was underestimated and that the mean size became veridical. However, the PSEs of the numerosity task increased as the mean size increased, $F(2, 6) = 5.174$, $p = 0.049$, indicating overestimations of numerosity. Therefore, the mean size was underestimated when numerosity was overestimated, and vice versa.

Interrelationship among the three magnitudes: Correlation analyses

To investigate the interrelationships among the estimations of the three magnitudes, we conducted a correlation analysis. In order to determine the correlations among the various estimations, we converted presented magnitudes into perceived magnitudes by multiplying the PSEs by the presented magnitudes. Since the PSEs that we measured were defined as the magnitude of difference (%) between the standard and test displays, the perceived magnitude can be acquired by multiplying the actual magnitude presented in the display by the observers' bias (i.e., PSEs). The correlations between perceived magnitudes showed that there were statistically significant correlations for all pairs of magnitudes. Estimations of summed area were positively

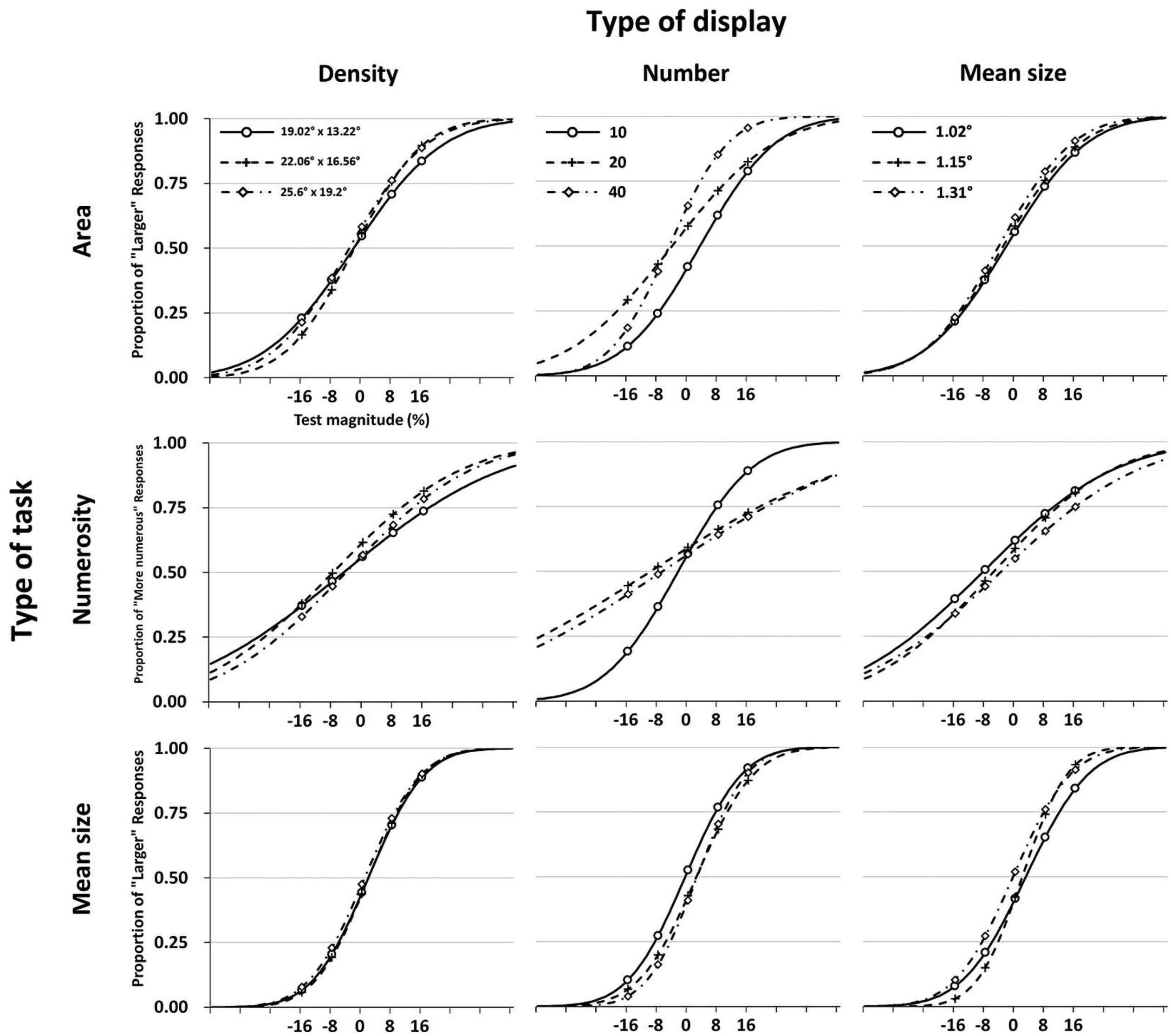

Figure 2. Psychometric functions of a representative observer in relation to three different tasks and three independent variables.

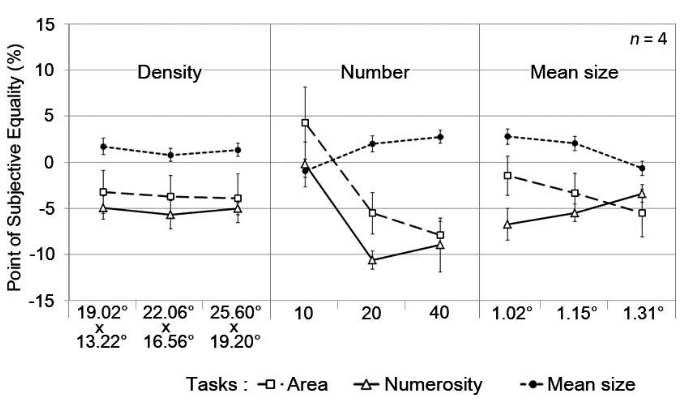

Figure 3. The PSEs of each task depending on three independent variables. Each column indicates different ways of generating displays, and each task was described as a separate line within each column.

correlated with estimations of both numerosity, mean $r = 0.97$, $t(3) = 222.564$, $p < 0.001$, and mean size, mean $r = 0.39$, $t(3) = 9.011$, $p = 0.003$. Numerosity estimations also showed a significant positive correlation with mean size estimations, mean $r = 0.23$, $t(3) = 7.592$, $p = 0.005$. Thus, all three magnitudes were related to each other, consistent with the generalized magnitude system hypothesis (Walsh, 2003).

Estimation sensitivity: Just-noticeable difference analyses

Figure 4 shows the just-noticeable difference (JND) for each condition. We first calculated the difference between the 75% threshold and PSE and then the difference between the 25% threshold and PSE. We then defined JND as the average of the two differences. A

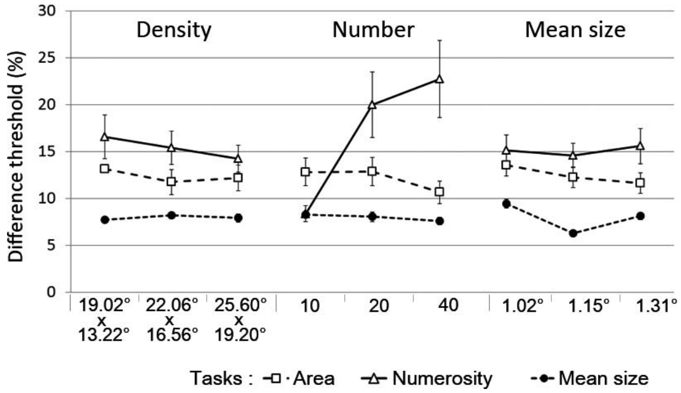

Figure 4. The JNDs for each task depending on three independent variables. Each column indicates different ways of generating displays, and each task was described as a separate line within each column.

smaller JND indicates greater sensitivity of a specific discrimination. First, we found a main effect of the tasks, $F(2, 6) = 11.655$, $p = 0.009$. The JND for the summed area task was significantly larger than that of the mean size task, $F(1, 3) = 30.041$, $p = 0.012$, and significantly smaller than that of the numerosity task, $F(1, 3) = 17.473$, $p = 0.025$, indicating that the ability to discriminate mean size was better than for discriminating summed area and numerosity. Second, as in the PSE analysis, the JNDs for the three tasks were unaffected by density variation, $F(2, 6) = 1.341$, $p = 0.330$ for summed area; $F(2, 6) = 2.375$, $p = 0.174$ for numerosity; $F(2, 6) = 0.880$, $p = 0.462$ for mean size. Third, regarding the effect of numerosity, while the JNDs for the summed area task were unaffected, $F(2, 6) = 2.317$, $p = 0.180$, the JNDs for the numerosity task significantly increased as the numerosity of the display increased, $F(2, 6) = 13.868$, $p = 0.006$. In contrast, the JNDs for mean size task decreased significantly as the numerosity of the display increased, $F(2, 6) = 5.463$, $p = 0.045$).[2] Similar patterns were found when the mean size varied. The JNDs for the summed area and mean size tasks significantly decreased as the mean size increased, $F(2, 6) = 12.208$, $p = 0.008$ for summed area; $F(2, 6) = 34.201$, $p = 0.001$ for mean size. These results indicate that the estimations of summed area and mean size improve when the mean sizes of the stimuli become larger. However, the JNDs for the numerosity task did not depend on mean size variation, $F(2, 6) = 1.668$, $p = 0.265$.

Percepts of each of the three properties predicted by the other two properties: Modeling approach

The analyses of the PSEs and JNDs suggest that the ability of observers to estimate mean size is more accurate and precise than their ability to judge summed area and numerosity. The PSEs of numerosity judgments were negatively biased, whereas those of mean size judgments did not demonstrate bias. Moreover, JNDs for mean size judgments were significantly smaller than those for the other judgments. Thus, it is unlikely that mean size is derived from the ratio of summed area to numerosity.

To further test whether mean size is derived from summed area and numerosity, we compared the perceived mean size, expressed by the measured *p*(greater)—probabilities of responding that a test display is greater than a standard display—from the dataset, to the mean size predicted by the other quantities (see Figure 5). The predicted performance in the mean size task was computed with a resampling method. For both the summed area and number tasks, *p*(greater) at five test levels were fitted to a cumulative Gaussian distribution (Figure 5a) so as to estimate the distribution (μ_{task},σ_{task}; Figure 5b) of observers' noisy percept. A predicted mean size was calculated by two random values, which was drawn independently from one from each distribution (Figure 5b for summed area and Figure 5d for numerosity). The ratio of these (the value from summed-area distributions divided by one from numerosity distribution) was used as a predicted mean size perception (Figure 5f). Repeating this process 1,000,000 times allowed us to reconstruct a predicted distribution (Figure 5g) of mean size perception, and *p*(greater) for the mean size judgment at five test levels could be predicted with the distribution.

In the same manner, we compared the predicted numerosity by the other quantities to the perceived numerosity and the predicted summed area by the other quantities to the perceived summed area. Specifically, the ability to estimate the summed area was predicted by multiplying mean size by numerosity and the ability to estimate numerosity was predicted by taking the ratio of summed area to mean size.

Figure 6 shows the measured results against the predicted results for each task. Three density levels were collapsed because it did not significantly influence observers' PSEs and JNDs. Figure 6a shows (1) observers' measured performance in the area task against (2) predicted performance for the area task computed by performance of the mean size and numerosity tasks. There were three levels of numerosity, three levels of mean size, and five levels of comparison stimuli for each observer. Thus, each graph in Figure 6 contains 180 ($3 \times 3 \times 5 \times 4$) data points. If observers inferred the mean size of stimuli from the ratio of total area to numerosity, the slope of the regression should be close to 1 (predicted performance = measured performance). In contrast, if observers perceived the mean size of stimuli directly, the slope would not necessarily be 1. In the area task, the slope of the regressed line was 1.145 (Figure 6a). To quantita-

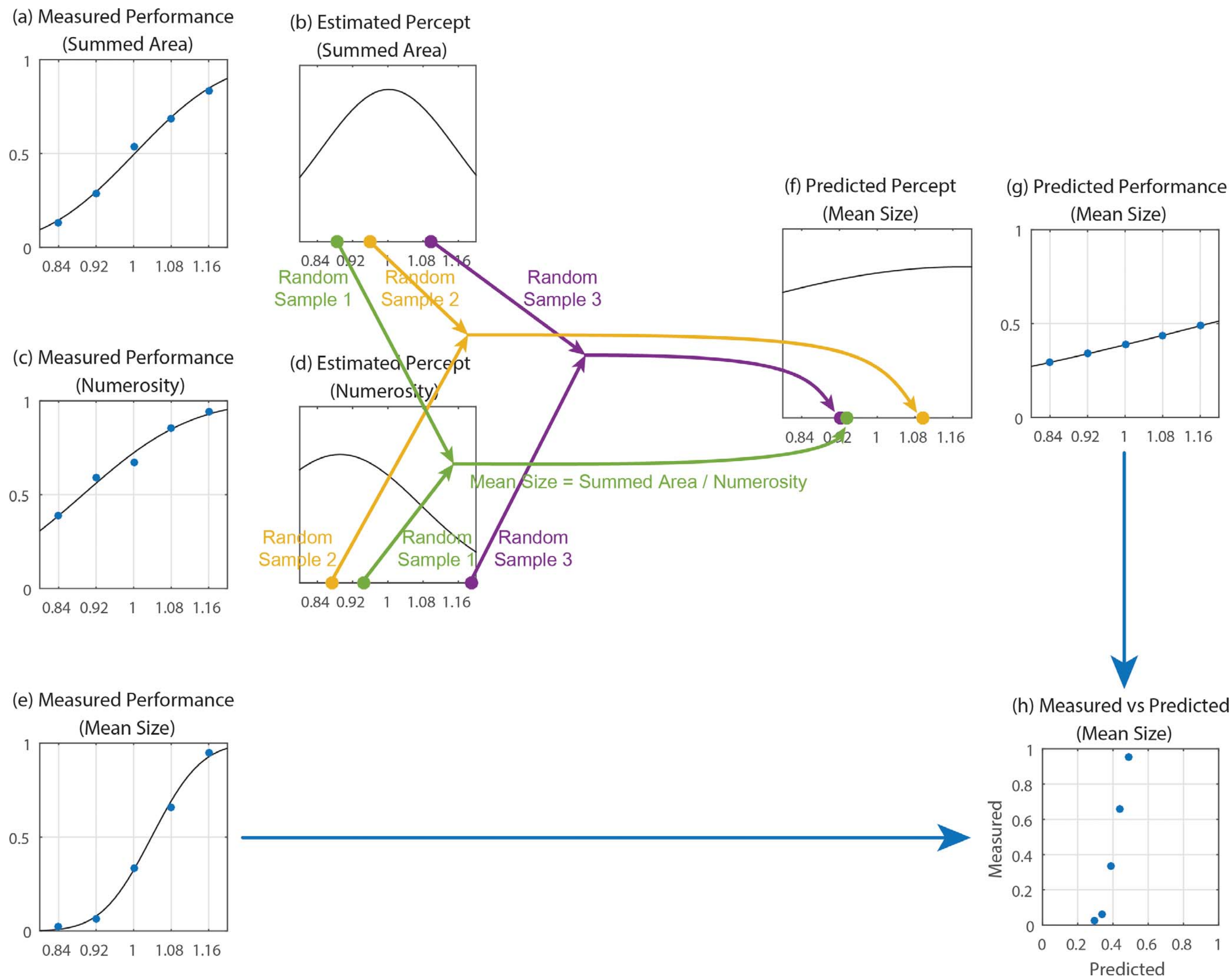

Figure 5. Graphical description of the quantitative analysis of performance-limiting noise. The units on the *x*-axis were transformed such that 0% difference between the standard and test displays became 1 to prevent dividing by 0.

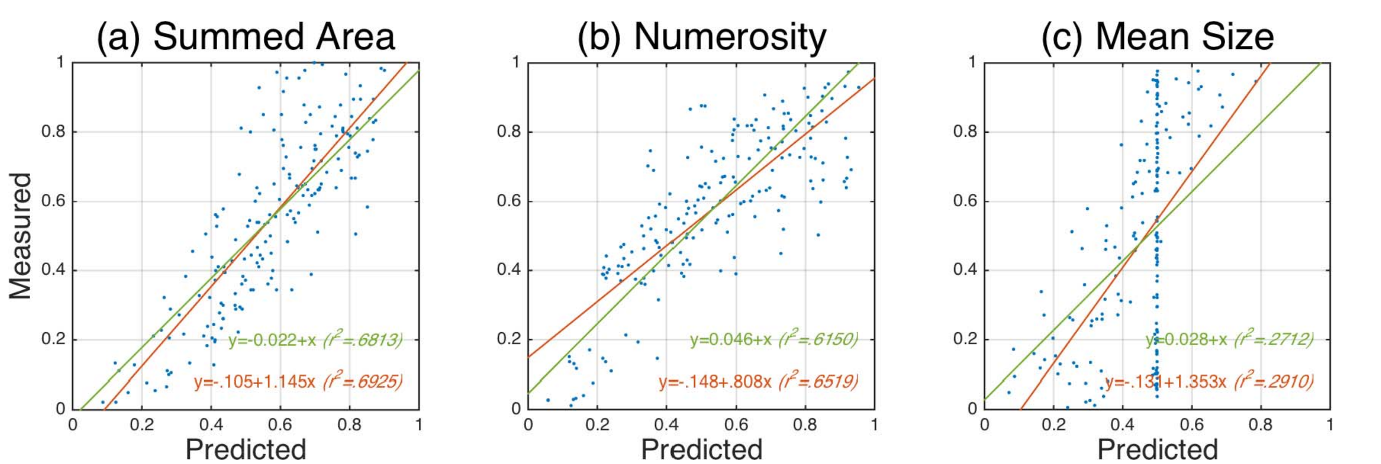

Figure 6. Measured versus predicted performance for each task. Green line indicates the slope of the reduced model and red line indicates the slope of the full model.

tively evaluate whether the slope was significantly different from 1, we compared the goodness-of-fit for the two regression models: the full model with a free slope parameter and the reduced model with the fixed slope parameter of 1. An F-test for nested models was used for statistical comparison (Wannacott & Wannacott, 1981). For the two models with k_{full} and $k_{reduced}$ parameters, the F statistic is defined as

$$F(df_1, df_2) = \frac{(r^2_{\text{full}} - r^2_{\text{reduced}})/df_1}{(1 - r^2_{\text{full}})/df_2} \quad (1)$$

where $df_1 = k_{\text{full}} - k_{\text{reduced}}$, and $df_2 = N - k_{\text{full}}$; N is the number of data points. For the area task, the slope was not significantly different from 1, $F(1, 178) = 0.695$, $p = 0.406$ (Figure 6a), suggesting that performance of the area task was well explained by those of the mean size and numerosity tasks. For the numerosity task, the slope (0.808) of the full model was significantly different from 1, $F(1, 178) = 18.869$, $p < 0.001$ (Figure 6b), suggesting that the numerosity task was not predicted well by the mean size and area tasks. For the mean size task, the slope (1.353) of the full model was significantly different from 1, $F(1, 178) = 4.971$, $p = 0.03$ (Figure 6c). These results suggest that performance in the mean size task was not predicted by the ratio of summed area and numerosity.

It should be noted that, in Figure 6c for the mean size task, some measured p(greater) were widely spread from 0 to 1, whereas corresponding predicted p(greater) were centered around 0.5. These occurred only when observers with significantly negative PSEs (Figure 3) and larger numerosity JNDs (Figure 4) judged mean sizes of 20 or 40 items. Those observers often extremely underestimated the numerosity of a display. If visual magnitudes (mean size, numerosity, and summed area) are perceived separately, perceived mean size should not be affected by such underestimation of numerosity. In contrast, if mean size was derived from the other magnitudes, a few extreme underestimations would yield large variation of mean size percept and, thus, lead observers to have very low discriminability of mean sizes. Our analysis shows that observers' actual performance was different from the prediction of the ratio model (data points aligned on the horizontal center in Figure 6c). In sum, the modeling results suggest that observers perceived the summed area using the other properties, but not numerosity and mean size.

Discussion

We investigated whether the visual system computes various magnitudes of a visual display based on other magnitudes. We found that the ability of observers to estimate mean size was better than their ability to estimate both summed area and numerosity. Specifically, mean size judgments did not demonstrate any bias, and had more sensitive discrimination thresholds than the others. The quantitative analysis of the performance-limiting noise showed that perceived summed area was predicted well by the other perceptual quantities, whereas both perceived mean size and numerosity were not predicted by the other perceptual quantities. Thus, perceived summed area is likely derived from multiplying perceived numerosity by perceived mean size. However, perceived numerosity is not likely to be derived from the ratio of perceived summed area to perceived mean size and perceived mean size is not likely to be derived from the ratio of perceived summed area to perceived numerosity.

Representing mean size is advantageous because the noise inherent in individual representations can be canceled out with the representation of mean size (Alvarez, 2011). We found that mean size discrimination thresholds became lower as the number of items increased. This is consistent with the prediction of a simple pooling model (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), given that the JNDs (8.3%, 8.1%, and 7.6%) in our study decreased significantly with increasing set size, almost proportionally to the square root of the number of elements. Allik et al. (2013) also found similar results, when the set size of 1 was not included in their analyses.

Our results also suggest that the visual system uses a distinct mechanism to compute numerosity, consistent with previous studies (Anobile, Cicchini, & Burr, 2014; Burr & Ross, 2008a; Liu, Zhang, Li, Zhao, & Tang, 2015; Liu, Zhang, Zhao, Liu, & Li, 2013; Ross & Burr, 2010). If the visual system computes and represents both mean size and numerosity, it does not necessarily have to represent summed area separately. Our results support this claim by showing that perceived summed area was predicted by multiplying perceived mean size by perceived numerosity. However, we should acknowledge that this evidence is indirect and the exact mechanism of computing numerosity has not been provided (Raphael & Morgan, in press). Indeed, many studies have suggested that perceived numerosity is based on the density of a display (Dakin et al., 2011; Durgin, 1995, 2008). When the size of a to-be-counted object is relatively small, the density of objects may determine the perceived numerosity of a display because the entire display of objects can be easily integrated into a texture. Previous studies (Dakin et al., 2011: 0.06°; Durgin, 1995: 0.04°; Durgin, 2008: 0.04°) that have found the effect of density on perceived numerosity had indeed smaller sizes than other studies that have found no effect (Burr & Ross, 2008a: 0.3°; Ross & Burr, 2010: 0.24°). Consistent with this explanation, we found no effect of density on perceived numerosity and our sizes (larger than 0.67°) were relatively larger than the previous studies. Thus, the exact mechanism of perceived numerosity can vary depending on the sizes of to-be-counted objects.

Because the visual world overwhelms us with a tremendous amount of redundant information (Kersten, 1987), the visual system reduces the redundancy using an efficient code of summary representation (Chandler & Field, 2007; Field, 1989). Thus, it is not efficient to represent all visual quantities, given the high correlation among them (Walsh, 2003). Consistent with this claim, our results suggest that the visual system computes summed area based on the multiplication of mean size and numerosity. Given high correlation ($r = 0.97$) between perceived summed area and perceived numerosity in our study, it is not efficient to represent both summed area and numerosity. Rather, the visual system should represent one property and use it to infer the other property. Our results suggest that the visual system represents numerosity and use it to infer summed area. If this is the case, observers should have difficulty in judging summed area when the number of items in a display does not covary with summed area. Indeed, incongruent numerosity interfered with judgments of summed area and incongruent summed area interfered with judgments of numerosity (Hurewitz, Gelman, & Schnitzer, 2006). Halberda, Sires, and Feigenson (2006) also found that errors of number estimation significantly increased when total area of to-be-counted objects did not covary with the number of objects. These results suggest that summed area and numerosity rely on a common metric.

Different magnitudes help one another to form a better representation in a given situation (Jacob et al., 2012). Jacob et al. (2012) proposed that the analog codes for the numerator and denominator and the representation of their proportions complement each other. Moreover, several previous studies (Ischebeck, Schocke, & Delazer, 2009; Jacob & Nieder, 2009a; 2009b) have suggested that proportions are represented in the intraparietal sulcus, the region of the brain commonly activated by various magnitude estimations (Pinel et al., 2004). Thus, it is easier for the visual system to utilize various magnitudes to form a better representation depending on a given context. Given the highly correlated context of stimuli between summed area and numerosity in our displays, the visual system is likely to utilize mean size and numerosity to infer summed area.

In summary, we investigated whether the visual system derives perceived magnitudes of visual properties (summed area, numerosity, and mean size) using other perceptual quantities. Our results showed that mean size discrimination was more precise than the other types of discrimination and that mean size judgments demonstrated less bias. In addition, we found that perceived summed area was predicted well by the other properties, whereas perceived mean size and numerosity were not predicted by the other properties. Thus, the visual system does not derive both mean size and numerosity from the other perceptual quantities; it might derive summed area from the other perceptual quantities.

Keywords: perceived magnitude, ensemble perception, area, numerosity, mean size

Acknowledgments

This work was supported by a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MEST, No. 2011-0025005). For helpful comments and discussion about this manuscript, we thank Eunsam Shin.

Commercial relationships: none.
Corresponding author: Sang Chul Chong.
Email: scchong@yonsei.ac.kr.
Address: Department of Psychology, Yonsei University, Seodaemun-gu, Seoul, Korea.

Footnotes

[1] Note that the PSE in the condition with 10 stimuli was significantly larger than that in the condition with 20, $t(3) = 5.907$, $p = 0.010$, whereas it did not differ between the conditions with 20 and 40 stimuli, $t(3) = -0.423$, $p = 0.701$.

[2] Allik, Toom, Raidvee, Averin, and Kreegipuu (2013) did not find this trend. We think that this discrepancy is due to not considering averaging noise in their study. In their study, JNDs mostly increased from the set size of 1 to the set size of 2, indicating that averaging noise was involved. In fact, when we removed the set size of 1 and reanalyzed their data, the effect of set size became significant, consistent with our results.

References

Allik, J., Toom, M., Raidvee, A., Averin, K., & Kreegipuu, K. (2013). An almost general theory of mean size perception. *Vision Research*, *83*, 25–39.

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics*, *49*, 303–314.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*, 122–131.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science*, *25*, 265–270.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*, 157–162.

Bauer, B. (2009). The danger of trial-by-trial knowledge of results in perceptual averaging studies. *Attention, Perception, & Psychophysics*, *71*, 655–665.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and non-temporal working memory tasks. *Perception & Psychophysics*, *59*, 1118–1140.

Burgess, A., & Barlow, H. B. (1983). The precision of numerosity discrimination in arrays of random dots. *Vision Research*, *23*, 811–820.

Burr, D., & Ross, J. (2008a). A visual sense of number. *Current Biology*, *18*, 425–428.

Burr, D., & Ross, J. (2008b). Response: Visual number. *Current Biology*, *18*, 857–858.

Casini, L., & Macar, F. (1997). Effects of attention manipulation on perceived duration and intensity in the visual modality. *Memory & Cognition*, *25*, 812–818.

Chandler, D. M., & Field, D. J. (2007). Estimates of the information content and dimensionality of natural scenes from proximity distributions. *Journal of the Optical Society of America A*, *24*, 922–941.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*, 393–404.

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research*, *45*, 891–900.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences, USA*, *108*, 19552–19557.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neuroscience*, *21*, 355–361.

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 149–169.

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*, *18*, R855–R856.

Field, D. J. (1989). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, *4*, 2379–2394.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, *16*, 123–144.

Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science*, *17*, 572–576.

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences, USA*, *103*, 19599–19604.

Ischebeck, A., Schocke, M., & Delazer, M. (2009). The processing and representation of fractions within the brain. An fMRI investigation. *Neuroimage*, *47*, 403–413.

Jacob, S. N., & Nieder, A. (2009a). Notation-independent representation of fractions in the human parietal cortex. *Journal of Neuroscience*, *29*, 4652–4657.

Jacob, S. N., & Nieder, A. (2009b). Tuning to non-symbolic proportions in the human frontoparietal cortex. *European Journal of Neuroscience*, *30*, 1432–1442.

Jacob, S. N., Vallentin, D., & Nieder, A. (2012). Relating magnitudes: The brain's code for proportions. *Trends in Cognitive Sciences*, *16*, 157–166.

Kersten, D. (1987). Predictability and redundancy of natural images. *Journal of the Optical Society of America A*, *4*, 2395–2400.

Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, *38*, 317–327.

Liu, W., Zhang, Z. J., Li, B. C., Zhao, Y. J., & Tang, Y. (2015). Numerosity adaptation along the Y-Axis affects numerosity perception along the X-Axis: Does numerosity adaptation activate MNLs? *Attention, Perception, & Psychophysics*, *77*, 1358–1370.

Liu, W., Zhang, Z. J., Zhao, Y. J., Liu, Z. F., & Li, B. C. (2013). Effects of awareness on numerosity adaptation. *PloS One*, *8*(10), e77556.

Onoe, H., Komori, M., Onoe, K., Takechi, H., Tsukada, H., & Watanabe, Y. (2001). Cortical networks recruited for time perception: A monkey positron emission tomography (PET) study. *Neuroimage*, *13*, 37–45.

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Pinel, P., Piazza, M., Bihan, L. D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*, 983–993.

Raphael, S., & Morgan, M. J. (in press). The computation of relative numerosity, size and density. *Vision Research*.

Ross, J. (2003). Visual discrimination of number without counting. *Perception*, *32*, 867–970.

Ross, J., & Burr, D. (2010). Vision senses number directly. *Journal of Vision*, *10*(2):10, 1–8, doi:10.2.10. [PubMed] [Article]

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature*, *415*, 918–922.

Stein, J. (1989). The representation of egocentric space in the posterior parietal cortex. *Experimental Physiology*, *74*, 583–606.

Teghtsoonian, M. (1965). The judgment of size. *American Journal of Psychology*, *78*, 392–402.

Vallentin, D., & Nieder, A. (2008). Behavioral and prefrontal representation of spatial proportions in the monkey. *Current Biology*, *18*, 1420–1425.

Vallentin, D., & Nieder, A. (2010). Representations of visual proportions in the primate posterior parietal and prefrontal cortices. *European Journal of Neuroscience*, *32*, 1380–1387.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488.

Wannacott, T. H., & Wannacott, R. J. (1981). *Regression: A second course in statistics*. New York: Wiley.