FEATURE SPECIFICITY

IN VISUAL STATISTICAL SUMMARY PROCESSING

HARUN YÖRÜK

BOĞAZİÇİ UNIVERSITY

2018

FEATURE SPECIFICITY IN VISUAL STATISTICAL SUMMARY PROCESSING

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

by

Harun Yörük

Boğaziçi University

2018

DECLARATION OF ORIGINALITY

I, Harun Yörük, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature...

Date ......24.07.2018........................

ABSTRACT

Feature Specificity in Visual Statistical Summary Processing

People initially process object sets by averaging their features into statistical summary representations. Previous studies have shown that there is independence between summarizing of low and high-level visual information. There is also evidence showing that, processing capacity of multiple statistical summaries is limited for simultaneous averaging of same kind of visual information, but not for different kinds of visual information. The current thesis investigates whether statistical summary processing relies on a feature-*specific* or a feature-*general* mechanism, and whether there are capacity limitations to simultaneous averaging of different visual features. We asked participants to average on of the features in a set of lines that varied in size and orientation. The relevant feature was either the same throughout a block or mixed within the trials of a blocks. Even though first two experiments showed a positive relation between viewers' size and orientation averaging performances for mixed averaging conditions, with more controlled displays we repeatedly found that performances on two tasks were unrelated both for single and mixed conditions. Viewers' errors for size averaging were higher in mixed than single averaging conditions, however this difference disappeared with reduced task difficulty. Orientation averaging performances were similar in single and mixed conditions. Finally, viewers' performances on size and orientation averaging tasks were similar across 50, 100 and 200 milliseconds of encoding durations. Overall, results of this thesis suggested that there are independent feature-specific statistical summary mechanisms for size and orientation features.

ÖZET

Görsel İstatistiksel Özet İşlemede Özelliğe-Özgüllük

İnsanlar öncelikli olarak obje setlerini onların özelliklerini ortalayarak istatiksel özet temsil olarak işlerler. Önceki çalışmalar alt-düzey ve üst-düzey görsel bilgilerin özetlenmesi süreçlerinde bağımsızlık olduğunu göstermiştir. Ayrıca çoklu özet işleme kapasitesinin aynı türde bilgiler için kısıtlıyken, farklı türde bilgiler içeren özetlerde kısıtlı olmadığını gösteren çalışmalar bulunmaktadır. Bu tez özet temsillerin özelliğe-özgül mü yoksa özelliklere-genel mekanizmalar tarafından mı işlendiğini ve izleyiciler aynı obje setinden iki özelliği eş zamanlı olarak özetlerken çoklu özellikler-arası özet temsillerin işlenmesinde bir kapasite limiti olup olmadığını incelemektedir. Katılımcılardan uzunluk ve yön açısından varyasyona sahip obje setlerinin özelliklerinden birini ortalamalarını istedik. Ortalanacak özellik deney bloku süresince aynı ya da deney blokunun denemeleri arasında karışık olarak sunuldu. İlk iki deneyin sonuçları izleyicilerin uzunluk ve yön özetleme performanslarının tekli özetleme koşullarında ilişkili olduğunu göstermişse de, daha kontrollü sahnelerle bu görevlerdeki performansların hem tekli hem karışık özetleme koşullarında ilişkisiz olduğunu bulduk. İzleyicilerin uzunluk ortalama hataları karışık koşulda tek koşuldan daha yüksekken, bu fark daha kontrollü sahnelerde kaybolmuştur. Yön ortalama hataları ise tekli ve karışık ortalama koşullarında daima benzer bulundu. Son olarak izleyicilerin uzunluk ve yön ortalama görevlerindeki performanslarının 50, 100 ve 200 milisaniyelik kodlama sürelerinde benzer kaldığını bulduk. Genel olarak bu tezin sonuçları yön ve uzunluk özellikleri için bağımsız özelliğe-özgül istatistksel özet temsili mekanizmaları olabileceğini önermiştir.

ACKNOWLEDGMENTS

I would like to present my deepest gratitude to Dr. Aysecan Boduroglu and Dr. Esra Mungan. I am thankful for their endless support, encouragements, wisdom and guidance throughout my research and education.

I want to thank Dr. Ahu Gökçe for having served in my thesis committee, and for her valuable comments and help.

I would also like to thank all my professors, colleagues, and friends in the Psychology Department of Boğaziçi University.

I want to thank TÜBİTAK (Scientific and Technological Research Council of Turkey) for providing financial support.

I want to give my special thanks to Nehir for making it all more enjoyable.

Lastly, I want to give my most special thanks to my mom and dad for believing in me and providing their endless support.

DEDICATION

To the memory of my beloved grandmother

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

People can extract average information from object sets by visual statistical summary processing (also called as ensemble processing) (for a review, see Alvarez, 2011). Previous research has shown that viewers can efficiently summarize various low-level features such as size (Ariely, 2001), color (Maule, Witzel & Franklin, 2014), brightness (Bauer, 2009), orientation (Dakin, 2001; Parkes, Lund, Angelucci, Solomon & Morgan, 2001), position (Alvarez & Oliva, 2008; Mutluturk & Boduroglu, 2016), as well as higher-level features like facial identity (de Fockert & Wolfenstein, 2009) and facial emotion (Haberman & Whitney, 2007). Viewers do not only extract summaries of static visual displays, they can also summarize sequentially presented visual (Albrecht & Scholl, 2010; Hubert-Wallander & Boynton, 2015) and auditory inputs (Albrecht, Scholl, & Chun, 2012; Piazza, Sweeny, Wessel, Silver & Whitney, 2013); they can also represent variance (Morgan, Chubb & Solomon, 2008; Semizer & Boduroglu, 2016) and numerosity information (Utochkin & Vostrikov, 2017). Despite demonstrations of statistical summarizing across various domains, interestingly, the mechanism underlying this ability is not clearly described. This research specifically investigates visual summarizing abilities to determine whether there is a domain general neural mechanism or whether there are multiple domain-specific neural mechanisms underlying this summarization ability.

Haberman, Brady and Alvarez (2015) proposed two possible mechanisms for how viewers extract visual summary information. One possibility is that there is a domain-general, central statistical summary processor in the visual system that is responsible for averaging all types of information. The alternative possibility is that there are multi-level, domain-specific mechanisms for summarizing different visual properties, possibly at different cortical levels. While the former view would necessitate performance in summarizing information in one domain to be correlated with summarizing performance in other domains, the latter mechanism would allow for independence of performance at summarizing different types of visual properties.

To determine whether there is a domain-general or domain-specific visual summarizing mechanism, Haberman, Brady and Alvarez (2015) compared summarizing performance across various visual properties. In their first experiment, they found that there was no correlation between orientation averaging and facial averaging tasks suggesting that statistical summary representations of high-level and low-level features may be governed by separate mechanisms. Subsequent experiments compared performance across various high-level (e.g. facial identity and facial emotion) and low-level (e.g. orientation and color) visual features. When viewers completed either two low-level or two high-level tasks, then there was a significant correlation between the two performances. On the other hand, when one task was from a low-level and the other one was from a high-level domain, there was no correlation between the two tasks. Thus, they concluded that statistical summary processing is not a uniform process, and that there are at least two separate and independent domain-specific summarizing mechanisms, for low-level and high-level visual information, specifically.

These findings were in line with the claims made by Whitney, Haberman and Sweeny (2014) who argued that there might be multilevel processing mechanisms in the ventral and dorsal pathways for summarizing information rather than a single cortical area responsible for summarizing all types of visual properties. They claimed that orientation, color, and brightness features might be summarized in early cortical stages, while motion and position information and size and shape features be summarized by separate mechanisms in the dorsal and ventral pathways, respectively. In addition, they argued that complex, higher-level face and biological motion summaries may be processed later, after the convergence of ventral and dorsal pathways While Whitney and colleagues outlined these possibilities, they did not provide empirical support for these claims; but recent work from our lab has provided indirect evidence that there may be more than one visual summary mechanisms. In a series of experiments, Uner and colleagues (2014) investigated how visual and spatial features were summarized.[1] Viewers completed two tasks, in one, they provided the mean length of a set of lines and in the other, and they estimated the centroid of a number of items. There was no relationship between how accurately people summarized line lengths and identified centroids. In a separate series of experiments, Yildirim and Boduroglu (2015) investigated if viewers can extract statistical summary representations of multiple features from the same set of objects. Participants studied displays that consisted of heterogeneously sized circles. Viewers then reported either the centroid (spatial summary) or the center-of-mass (the weighted spatial center, which required viewers to combine size and spatial information). Viewers were equally accurate at retrieving both types of summaries suggesting that they could integrate visual and spatial properties of the studied set to produce a combined summary. More critically, they directed viewers to attend to size and asked either the location of the smallest/largest object or the centroid within

[1] The goal of this experiment was to compare action video-game players with strategy video-game players on centroid (mean position) and mean size estimation tasks. In centroid estimation task, they saw 7 squares scattered around the display and asked to report centroid of the squares. In mean size estimation task, they saw 9 lines with various lengths and asked to report the mean size of the lines. Uner and colleagues (2014) found that there was no significant relation between the performances across these two domains in the overall sample. (r =.161, p =.161). Only for the AVGP group, there was a small yet significant relationship between the two domains, suggesting that these summarizing abilities may be benefitting from video-game related enhancements in visual processing skills (r = .42, p =.05).

the same block to see if there is a bias towards the center-of-mass in centroid responses because of a possible binding process. They found that attending to the size did not result in a bias towards the center-of-mass in centroid responses, compared to single centroid block. This result suggested that visual and spatial summaries may not be obligatorily bound (Yildirim & Boduroglu, 2015). Altogether these findings demonstrated that while viewers can summarize size and space information in a simultaneous and interactive manner (as revealed by the center-of-mass results), they can also independently summarize these features given task demands. These two studies suggest that re there could be some level of independence between the summarizing of low-level features in the visual and spatial domains.

In sum, the empirical research so far, has suggested that there may be separate mechanisms responsible for the summarization of lower and higher-level complex visual features, as well as visual and spatial features. This current research further tested the domain-specificity argument by comparing how viewers summarize features that belong to the same domain (e.g. size and orientation features of visual domain). We expected independence for statistical summary processing among low-level visual features.

CHAPTER 2

THE PRESENT STUDY

To test this general prediction, we asked participants to summarize two sets of features, across different experiments. In these experiments there were blocks of trials in which viewers summarized only one of the critical features (single-task blocks); there were also a separate set of mixed feature trials where viewers were asked to report the summary of either one of the features, identified by a cue following display offset (mixed-task blocks).

As summarized above, a feature-general averaging mechanism would necessitate a dependence between the errors across the two features, thus a correlation between the performances. In addition, if there is a feature-general mechanism, then in mixed-task trials, there may be more interference between summarization of features, resulting in higher error compared to single-task trials. In contrast to this alternative, a feature-specific mechanism would allow the summarization errors to be independent, and even in mixed blocks interference between summarization of different features would be unlikely, resulting in similar levels of errors in single and mixed task blocks.

2.1 Experiment 1A

In Experiment 1A, we tested how averaging performance in one visual feature predicted performance in the other feature by focusing on mean size and mean orientation estimates. We presented people with displays that consisted of lines with different lengths, displayed at various orientations and they reported either the mean length or the mean orientation of these lines. While the errors in these two conditions may be related because both features are visual in nature, the literature suggests that it is more likely that these low-level visual features may be still independently averaged. In particular, orientation may be summarized within the early cortical regions; size information may be summarized along the ventral stream (Whitney, Haberman & Sweeny, 2014). Therefore, we expected participants' errors in size and orientation averaging tasks to be unrelated, with independent feature-specific mechanisms.

2.1.1 Method

2.1.1.1 Participants

Twenty-six Bogazici University undergraduate students participated in the experiment in return of course credit for their participation. We excluded data of two participants because they did not generate a response for approximately half the trials, and an additional participant was more than 3 standard deviations away from the group mean on reaction time. Therefore, we conducted the analyses on the data from 23 participants.

2.1.1.2 Materials and stimuli

The experiment was programmed in E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA) and was run on a 17' monitor with screen resolution set to 1024x768 pixels (32x24 cm). Participants were sitting approximately 57 centimeters away from the computer screen. From that viewing distance, 1 cm was equal to 32 pixels and 1º visual angle.

Each trial (see Figure 1) began with a green fixation cross, presented for 1500 milliseconds, which turned red for 500 milliseconds to indicate the beginning of a trial. We instructed the participants to fixate on the cross when it turned into red. The object displays that consisted of 12 randomly oriented lines of unique lengths were presented for 200 milliseconds. Then, participants heard an auditory cue indicating which feature of the object sets they should average (either Orientation or Size), and were presented with a response probe. Participants had to adjust the length of the line for the size averaging trials, or rotated the orientation angle of the line for orientation averaging trials by the left button of the mouse, and they completed the trial by clicking to the right button of the mouse.

In each display, the orientation of the lines were randomly determined from a normal distribution ranging from 1 to 180 degrees. Line lengths were randomly selected from a range of 24 and 152 pixels (.75° – 4.75° of visual angle). No two lines were of the same length or orientation as the set mean and there were no repetitions of length or degree within sets. Lines were positioned on a 5x4 invisible grid (760x608 pixels), and 3 locations were randomly chosen from each row. We slided middle two rows of the grid by ±16 pixels in order to prevent endpoints of lines touching the center of the display. Lines were drawn in white on a gray background.

On the response screen, either a blue or green response probe was presented to distinguish the two types of trials. We determined a random value between 21 and 27 pixels, and, that value added either the half of the mean length of the set of lines or subtracted from the mean length to determine the length of response probe for the size adjustment task. In the orientation rotation task the angle of the probe line was determined by either adding to or subtracting a random number between 21 and 27 degrees from the mean angle of the set of lines.

2.1.1.3 Procedure

Before the actual experiment trials, participants completed a training session. To demonstrate viewers what average size of a set of lines looks like, we presented example displays, which consisted of 2,3,4,6,8,10, and 12 lines, sequentially. Each display was followed by another display that had an additional red line that indicated the mean length of the lines. After that, we gave 10 trials of the size averaging task for practice. We presented the displays for 1000 milliseconds in the training trials. After their response, we gave participants a feedback by showing the correct response line in red. We followed the same procedure for orientation averaging task. After the training, participants received a 16 trial practice block, which had the same procedure as the actual experiment. In the actual experiment, there were 150 size and 150 orientation averaging trials in total, divided in 3 experimental blocks. We presented two tasks in an intermixed fashion.

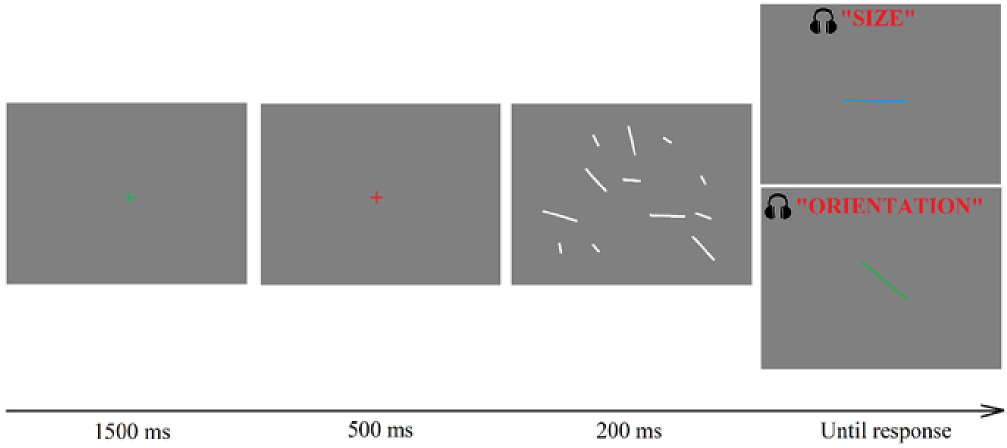

Figure 1. Example of a trial

2.1.2 Results

For the size averaging tasks, we calculated participants' performance by calculating the absolute pixel difference between the mean line size and the response. For the orientation averaging task, we calculated the absolute acute angle difference between the mean orientation of the lines and the orientation of the response line. The normality assumption was not satisfied for mean size estimation errors (Shapiro-Wilk $p < .05$); therefore, we conducted nonparametric Spearman's correlation analysis. Surprisingly, there was a significant positive correlation between mean size (M = 25.94, SD = 10.94) and mean orientation estimation errors (M = 32.42, SD = 8.57), r = .436, p = .043.

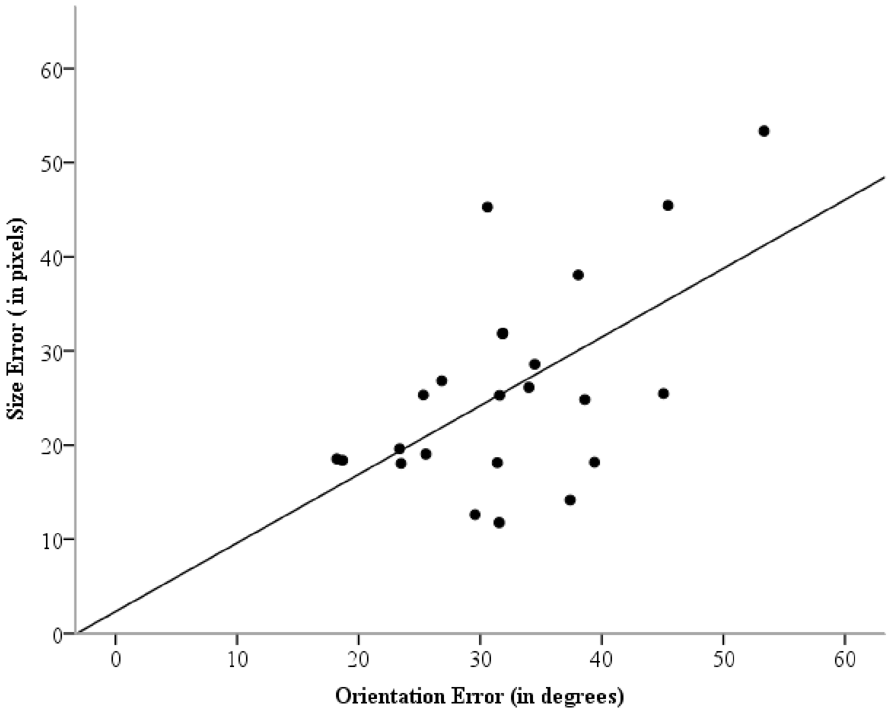

Figure 2. Correlation between size and orientation averaging performances.

We wanted to determine whether this unexpected pattern of relationship, was partly driven by task difficulty. If viewers were responding randomly in both tasks, then their performance may have seemed related, not because of a feature-general mechanism but because of random performance. Thus, we wanted to check whether viewers were randomly responding in the mean orientation and mean length trials.

We ran two simulations depicting random-like behavior in the orientation task. In the first simulation, all responses were assumed to come from the range of correct mean orientations of the displays (55 and 136 degrees). In the second simulation, we assumed that all responses to come from all possible range of responses (1 to 180 degrees) to simulate a total random observer performance. Univariate ANOVA conducted to see whether the experimental data was closer to Simulation 1 or 2 revealed that experimental errors were significantly higher than data in Simulation 1. However, experimental errors were significantly lower than Simulation 2 (*p*s < .001); and experimental data (*M* = 31.81, 95% CI [29.62, 34.01]) was closer to Simulation 1 (*M* = 24.42, 95% CI [22.22, 26.62]) than Simulation 2 (*M* = 45.42, 29% CI [43.10, 47.48]). (See Figure 3). This pattern of results suggests that viewers were not fully sensitive to task parameters, and yet were not performing the orientation task in a fully random fashion, either.

We ran two similar simulations for the mean size task. In the first simulation, the responses were always chosen from the range of the possible mean sizes depicted in the sets (61 to 110 pixels); in the second simulation a total random observer's performance was generated by picking responses between the minimum and maximum individual line lengths presented in the displays (24 to 152 pixels). The comparison of the experimental data to the simulation data for the mean orientation task revealed a very similar pattern to that reported for the orientation task. A

Univariate ANOVA revealed that, experimental errors were significantly higher than that in Simulation 1, but lower than Simulation 2 ($ps < .001$); and experimental data ($M = 23.11$, 95% CI [13.36, 17.85]) was closer to Simulation 1 ($M = 15.61$, 95% CI [13.36, 17.85]) than Simulation 2 ($M = 33.98$, 95% CI [31.74, 36.22]).

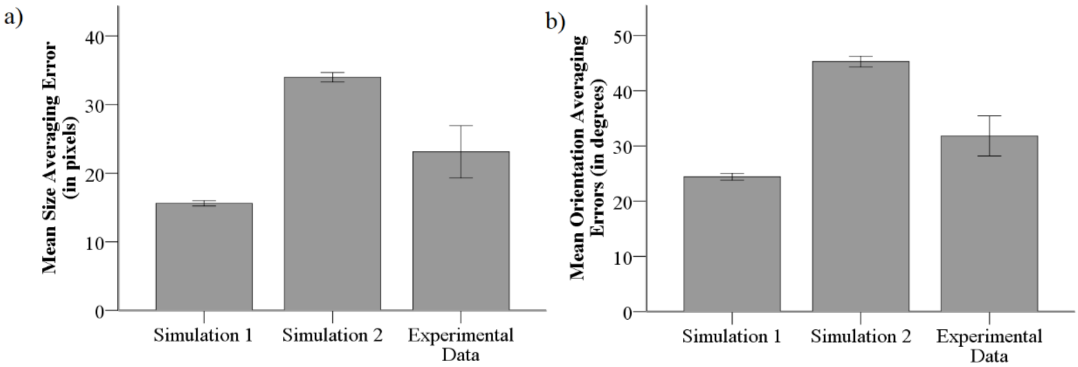

Figure 3. Random observer simulations for (a) size averaging and (b) orientation averaging tasks. Error bars indicate ±2 standard error of the mean.

2.1.3 Discussion

Contrary to our expectations, we found that there was a strong correlation between the size and orientation errors in Experiment 1A. This suggests that size and orientation features may be averaged by the same feature-general mechanism. However, the fact that we only used a mixed block in this experiment, allows for an alternative explanation. Since participants did not know which feature they were going to be asked to report, optimum performance would have required viewers to summarize both features and then choose the right feature to report following the cue. Thus, the mixed-task design may have created an additional demand on the viewers to summarize both features concurrently. This demand may be partly responsible for the relatively poor performance; on both tasks, viewers responded in a somewhat random fashion and reported means outside the range of possible means.

This suggests that participants were not fully sensitive to task parameters and their responses may be reflecting a lower boundary, shaped by the attentional demands of the design. This interpretation is possible in light of debate surrounding the view on whether there are costs associated with the summarization of two separate features (for a debate on this issue see Attarha & Moore, 2015b and Emmanouil & Treisman, 2008). Emmanoil and Treisman, (2008) suggested that there was an attentional cost related to averaging multiple sets with different visual features. On the other hand, Attarha and Moore (2015b) suggested that attentional cost of processing multiple averaging was evident when sets varied on the same feature dimension but not when they varied on different feature dimensions.

2.2 Experiment 1B

The correlation we obtained in Experiment 1A may in part reflect individual differences in attentional capacity as opposed to individual differences in summarization ability. Therefore, to eliminate this confound, in Experiment 1B, in addition to mixed-trial blocks, we included single-trial blocks where viewers had to summarize only orientation or size information. If statistical summaries for size and orientation features are extracted by independent feature-specific mechanisms, then viewers should be able to extract summaries for both of these features simultaneously, we should not expect reduced performance in mixed condition compared to the single task condition. On the other hand, if both summaries are extracted by a domain-general mechanism for low-level visual features, then it is more likely that there would be costs associated with summarizing both size and orientation in the mixed- compared to the single-block trials.

Previous research has demonstrated that number of objects within a set does not impact accuracy of statistical summary representations, suggesting that extraction of these representations is an automatic process, and does not have the limitations of attentional capacity bottleneck (Ariely, 2001; Chong & Treisman, 2003; Chong & Treisman, 2005). Chong and Treisman (2005) even showed that there was no additional cost when viewers had to summarize two simultaneously presented sets on the same feature dimension (e.g. size). However, there is not yet a full consensus on how attention influences summarization of different features. While Emmanoil and Treisman, (2008) showed that, when the viewers had to summarize two sets on different dimensions (size vs. speed, size vs. orientation) simultaneously, there was an associated cost of dividing attention regarding accuracy of the summary representations. They used a precue-postcue paradigm to determine this. In those experiments, participants knew which feature to attend in precue trials before encoding stage, whereas they did not know in postcue trials. They found that there was a cost of accuracy in postcue trials by higher average estimation errors compared to precue trials. Using a similar method, Utochkin and Vostrikov (2017) found that there was no cost of extracting mean size and numerosity information from the same object sets, which suggested an independence between averaging processes of two different statistical properties.

Contrary to these previous findings, later research found that capacity for simultaneous statistical summary processing was limited when two sets were on same feature dimensions (e.g. both sets varied on size or orientation), whereas it was unlimited when they were on two different dimensions (one from size and one from orientation). Therefore, there was a cost of attention when the two averaged features were on the same domain, but not when they were different visual features. (Attarha, Moore & Vecera, 2014; Attarha & Moore, 2015a; Attarha & Moore, 2015b).

2.2.1 Method

2.2.1.1 Participants

Twenty-five Bogazici University undergraduate students participated in the experiment in return of course credit for their participation. We excluded data of two participants who had size averaging errors more than 3 standard deviations away from the group mean. Therefore, we conducted the analyses on the data from 23 participants.

2.2.1.2 Materials and procedure

Experiment setting, displays, and the task procedures were the same as in Experiment 1 except for the inclusion of single task blocks. The first two blocks were single task blocks with size and orientation averaging tasks counterbalanced across participants; the subsequent two blocks were mixed-trial blocks. There were 90 trials in each block, with 360 trials in total.

2.2.2 Results

A pairwise t-test showed that errors for size averaging was significantly lower in single (M = 18.46, SD = 6.15) than mixed task conditions (M = 23.86, SD = 12.03),

($t(22) = -2.53$, $p < .05$, Cohen's $d = .513$). However, there was no significant difference between single ($M = 33.85$, $SD = 7.88$) and mixed ($M = 32.43$, $SD = 8.85$) conditions for orientation averaging tasks ($t(22) = -.92$, $p = .370$, Cohen's $d = .169$) (Figure 4). According to Pearson's correlation analysis, the relationship between size and orientation errors were not significant in the single block conditions ($r = .108$, $p = .625$), but there was a marginally significant and positive relationship between size and orientation averaging errors in the mixed blocks conditions ($r = .361$, $p = .090$) (see Figure 5). The overall pattern suggests that the size averaging task may have been impacted by the attentional demands in the mixed block, while orientation averaging was not. The lack of a correlation between the errors in the two tasks in the single block, and the weak yet marginally significant correlation in the mixed block may reflect the additional attentional demands in the mixed block trials.

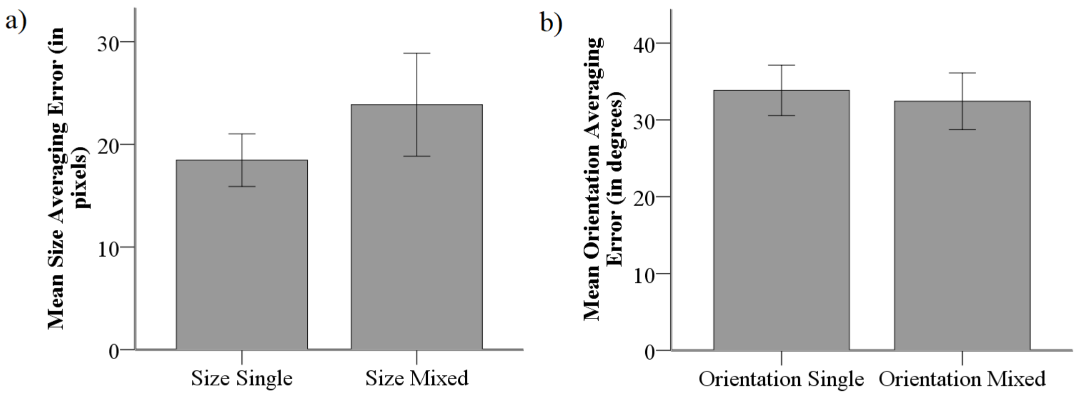

Figure 4. Mean (a) size and (b) orientation averaging errors across single and mixed conditions. Error bars indicate ±2 standard error of the mean.

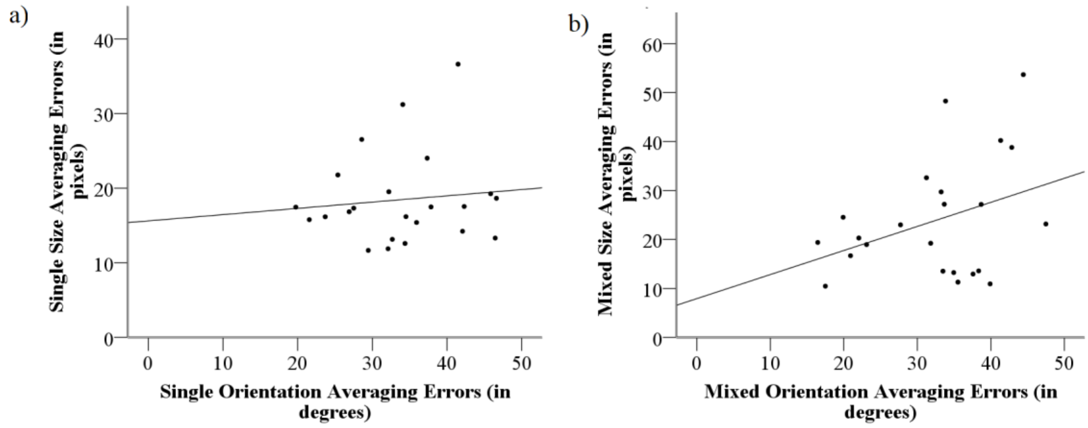

Figure 5. Correlations between (a) single and (b) mixed size and orientation averaging tasks.

To determine the impact of amodal attentional demands on performance, for each participant, we calculated an attentional cost score by subtracting their performance on the single conditions from their performances on the mixed conditions separately for the size and orientation averaging tasks (size-cost and orientation-cost, respectively). Pearson's correlation analyses showed that there was no significant relationship between the size-cost and orientation-cost. (r = .072, p = .744), suggesting that there was not a broad attentional limitation guiding performance. The size-cost did not correlate with performance in single size averaging block (r = .049, p = .826), but it did with the performance on size averaging trials in mixed conditions (r = .755, p < .001) suggesting that size averaging is influenced by attentional capacity. However, for orientation trials, there was no correlation between performance in the orientation-cost in neither single nor mixed blocks (r = -.084, p = .703 and r = -.166, p = .448, respectively).

In summary, we found that there was no correlation between size and orientation averaging performances, but only the performance on the length task was open to attentional influences. This pattern of findings is consistent with the idea that length and orientation features could be averaged independently. However, there remains an alternative explanation of the results. If participants were responding more randomly in the orientation task compared to the length task, random performance could have caused there to be no relationship between the single and mixed trials for these trials. To rule out this possibility we generated randomly simulated data (as in Experiment 1A) and demonstrated that people did not randomly respond in the orientation task. Overall, we found significant differences between simulated data, which assumed random responses, and experimental data suggested that participants were not performing in a random fashion in size and orientation averaging tasks. Therefore, we eliminated the possibility of random responses as an explanation of not finding any relation between size and orientation averaging errors in single and mixed blocks.[2]

2.2.3 Discussion

Experiment 1B showed that errors in size and orientation summaries were not correlated for the single task conditions, but were marginally correlated for the mixed

---

[2] The criteria used for the response ranges were identical with that reported in Expeirment 1A. For single size averaging tasks experimental errors were marginally significantly higher than data in Simulation 1 (p = .07), but lower than Simulation 2 (p < .001); and experimental data (M = 18.46, 95% CI [16.91, 20.01]) was closer to Simulation 1 (M = 15.91, 95% CI [14.35, 17.46]) than Simulation 2 (M = 33.02, 95% CI [31.47, 34.57]). For mixed size averaging tasks experimental errors were significantly higher than data in Simulation 1, but lower than Simulation 2 (ps < .001); and experimental data (M = 23.86, 95% CI [20.88, 26.84]) was slightly closer to Simulation 2 (M = 33.31, 95% CI [30.33, 36.29]). Than Simulation 1 (M = 14.64, 95% CI [11.66, 17.63]).
For single orientation averaging tasks experimental errors were significantly higher than data in Simulation 1, but lower than Simulation 2 (ps < .001); and experimental data (M = 33.85, 95% CI [31.72, 35.99]) was slightly closer to Simulation 1 (M = 24.46, 95% CI [22.33, 26.60]) than simulation 2 (M = 44.83, 95% CI [42.69, 46.96]). For mixed orientation averaging tasks experimental errors were significantly higher than data in Simulation 1, but lower than Simulation 2 (ps < .001); and experimental data (M = 32.43, 95% CI [30.19, 34.67]) was closer to Simulation 1 (M = 23.54, 95% CI [21.30, 25.78]) than Simulation 2 (M = 45.00, 95% CI [42.72, 47.20]).

task conditions. The lack of correlation in the single task conditions are consistent with feature-specific independent summary mechanisms. The marginal correlation in the mixed block may be in part due to the attentional demands due to the mixing of the two tasks.

An unexpected difference emerged between the size and orientation task. While there was no difference between the single and mixed condition for the orientation task, there was a difference between these two conditions for the size task.

There may be several possible explanations for this pattern. One reason why errors were similar in the single and mixed orientation trials may have to do with the difficulty of the displays and low levels of performance in the orientation task. If performance is approaching floor for the orientation task in both conditions, the single vs. mixed manipulation may not have worked. However, the comparison of the random response simulations and the actual data suggests that people were not responding in a totally random fashion. Nevertheless, the specific task parameters may have increased the task difficulty to the extent that the performance was approaching the lower bounds

A second possible explanation for the no difference in orientation errors in the single and mixed blocks may be related to the possibility that orientation averaging is a pre-attentive process, and thus it may precedes any attentional bottleneck. Myczek and Simons (2008) had indeed argued that early visual system primarily processes mean percept of orientation, color, and motion features, but not size to extract statistical summary representations. Therefore, orientation might not be open to attentional interference of multiple summary processing unlike size feature.

A third possible explanation is that size and orientation features can be processed at different levels of statistical summary representation mechanism as Attarha and Moore (2015b) suggested. It is possible that summary processing of orientation can be earlier in the hierarchy than size averaging. For example, orientation can be averaged early in the time frame of statistical summary processing, before size averaging can interfere. Whereas orientation might have interfered with averaging of size feature at earlier stages.

2.3 Experiment 1C

The goal of Experiment 1C was to determine whether task characteristics had an impact on the performance in the orientation-averaging task. To examine this issue, we systematically varied the range of the lines presented and tested whether error increased as the range of the orientation presented became larger.

2.3.1 Method

2.3.1.1 Participants

Nine Bogazici University undergraduate students participated in the experiment and we compensated them with course credit in return of their participation.

2.3.1.2 Materials and stimuli

In this experiment, we only had the orientation condition. All parameters were the same with the previous orientation averaging trials (single block) except that we manipulated line lengths, orientation ranges and controlled tilt direction. For each display, we chose the orientations randomly from 15, 30, 60, 90, 120 degrees ranges. On half of the trials line lengths were kept uniform (88 pixels), in the other half line lengths were chosen randomly from a 24 and 152 pixels range. Also, the direction where the orientation of the tilted lines were controlled as left, right, and even at both sides (For the displays with 120 degree ranges, which have both left and right tilted lines, we considered the mean direction as tiltedness direction.) Participants completed a total of 260 trials.

2.3.2 Results

Univariate ANOVA across displays showed that, the error in orientation averaging task significantly increased as the angle range of the lines in the displays increased ($F(4,260) = 49.32$, $p < .001$, $\eta_p^2 = .457$). Variance of the line sizes were also significantly influenced orientation averaging errors ($F(1,260) = 4.37$, $p < .05$, $\eta_p^2 = .018$) (Figure 6). Tiltedness of the lines had no significant effect on orientation averaging errors ($F(2,260) = .287$, $p = .751$, $\eta_p^2 = .002$). These findings suggest that the performance in Experiment 1B may have been negatively impacted by display characteristics, particularly line orientations.

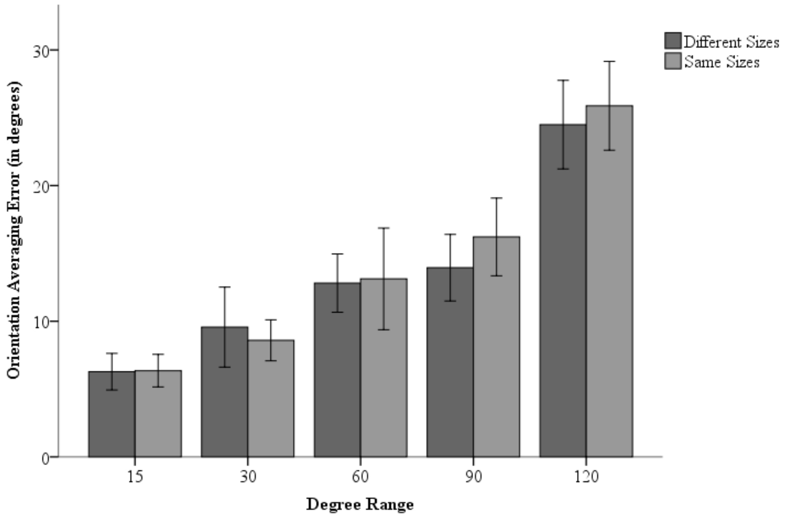

Figure 6. Mean orientation averaging errors across degree ranges and size variances. Error bars indicate ±2 standard error of the mean.

2.4 Experiment 2

Experiment 2 aims to rule out the possibility of random responses caused by difficulty of orientation averaging task, we therefore used 60 and 90 degrees as orientation ranges to keep the task difficulty at an optimum level. We expected orientation averaging errors to be lower than Experiment 1B, by the reduced task difficulty.

2.4.1 Method

2.4.1.1 Participants

Twenty-five Bogazici University undergraduate students participated to the experiment, in return for course credit. We excluded data of two participants because their size and orientation averaging errors (respectively) were more than 3 standard deviations away from the group mean, one participant who gave irrelevant responses, and one participant who was color-blind. Therefore, we conducted the analyses on the data of 21 participants.

2.4.1.2 Material, stimuli and procedure

The same procedure and design was used as in Experiment 1B except that the orientation ranges of the lines were selected from either a 60 or 90 degree range. Also, we counterbalanced the single and mixed block conditions across participants.

2.4.2 Results

Pairwise t-tests showed that, there was no difference in errors in single (*M* = 19.61, *SD* = 7.76) versus mixed size trials (*M* = 18.51, *SD* = 6.25), *t*(20) = .702, *p* = .140, Cohen's *d* = .154. There was also no difference between averaging errors for single (*M* = 15.22, *SD* = 3.73) and mixed orientation averaging tasks (*M* = 14.88, *SD* = 2.83), *t*(20) = .808, *p* = .429, Cohen's *d* = 0.102 (see Figure 7). In addition, there was no significant correlation between size and orientation averaging errors in neither the single nor the mixed blocks (*r* = -.076, p = .745 and (*r* = .-107, *p* = .644, respectively) according to Pearson's correlation analysis (see Figure 8). Finally, variability of line sizes did not influence orientation averaging errors.[3]

[3] We wanted to check if variability on line sizes have an impact on orientation averaging errors, because shorter lines in the sets could have been less influential by being less salient than longer lines. Also, tasks irrelevant features could be influential on statistical summary representations (Brady & Alvarez, 2011). We determined the variability of line sizes by calculating standard deviations of the sizes of 12 lines, and we divided them in 3 categories as low, medium, and high variance. There was

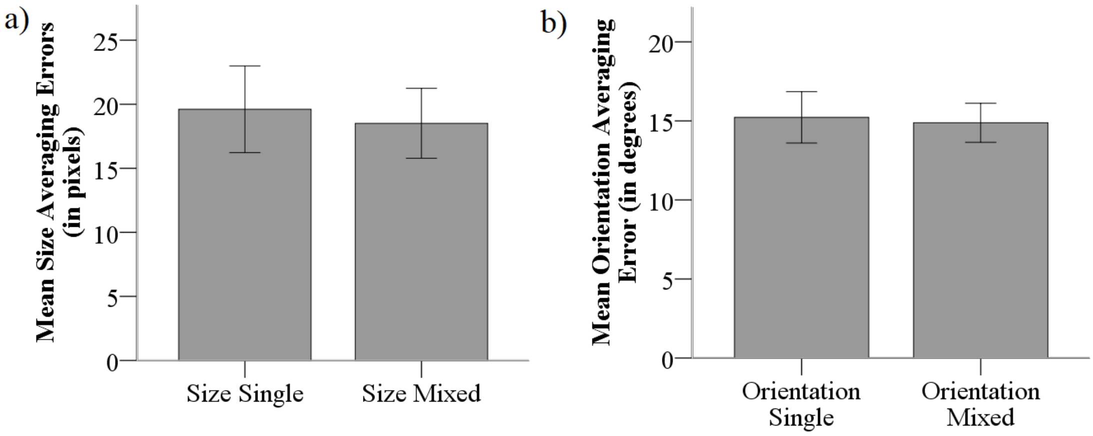

Figure 7. Mean (a) size and (b) orientation averaging errors across single and mixed conditions. Error bars indicate ±2 standard error of the mean.

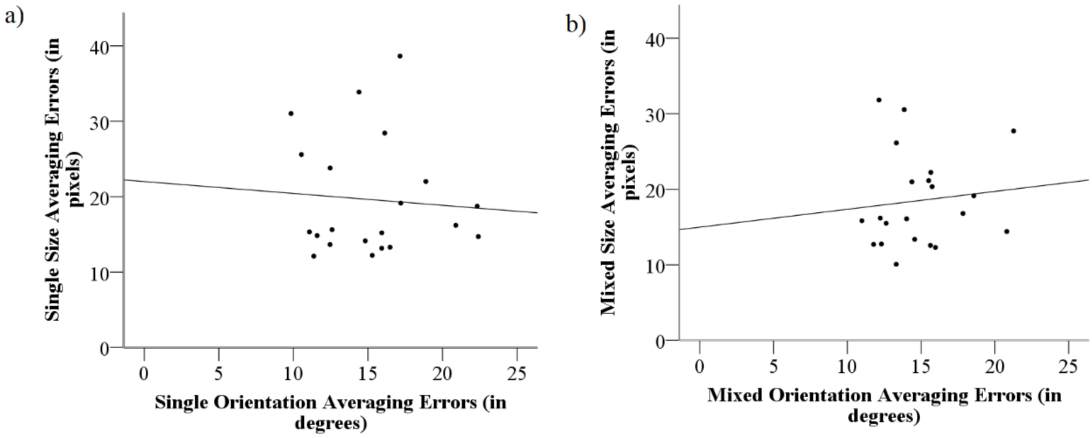

Figure 8. Correlations between (a) single and (b) mixed size and orientation averaging tasks.

2.4.3 Discussion

In Experiment 2 we demonstrated that in both the single and mixed conditions, there was no correlation between length and orientation averaging errors. Also, the levels

---

no significant effect of size variability on orientation averaging errors (F(2,180) = 2.816, p = .063, η_p^2 = .031).

of errors were similar in both the length and orientation task across the single and mixed conditions, suggesting that the uncontrolled display characteristics in Experiment 1B actually interfered with performance. Overall, these results provided us more evidence for independent feature-specific statistical summary processing mechanisms for size and orientation averaging tasks.

Our results seem to be in line with Attarha and Moore's (2015b) findings on capacity of between feature statistical summaries. They found that, there was a cost related with simultaneous processing of within-feature summaries (two sets that varied on size or orientation); however, there was no cost of between-feature summaries (one size and one orientation feature), suggesting size and orientation are summarized by separate independent mechanisms. On the other hand, our findings contradict Emmanouil and Treisman (2008)'s findings. Emmanouil and Treisman (2008) asked people to average one of the two spatially segregated sets which were varied on either size or orientation feature. They used a precue-postcue design. In precue trials participants were given a cue before the displays where they know which feature to attend before encoding, whereas in postcue trials participants received the cue after presentation of the displays. With that design, Emmanouil and Treisman (2008) found a cost associated with dividing attention between two sets in a statistical summary estimation task, when one of the sets varied in size and the other varied in orientation. However, there were some specific differences as well as similarities between our task and theirs. As similarity, participants were directed to attend one of the features before encoding stages in single block conditions of our study, and precue condition of Emmanouil and Treisman's study (2008); while participants were not given a specific feature to attend in mixed block of our study, and postcue condition of theirs until the encoding stage is completed. One of the differences was that, in our design we directed viewers to extract summaries from a single set that varied on two features; whereas Emmanouil and Treisman's (2008) task required viewers to extract summaries from two different sets spatially located on two sides of the screen. Therefore, viewers may have attended only one of the features by fixating one side of the display that contained the set with the relevant feature in precue condition. Whereas in our design even though participants knew which feature to attend in single block condition, the irrelevant features was still in their fixation area. Therefore, there could be an advantage in Emmanouil and Treisman's (2008) precue conditions, stemmed from fixating viewers' attention to the relevant set on a certain area, while the set with irrelevant feature is ignored. In relation with this, we know that, visual working memory could allocate more resources and provide higher resolution representations to the fixated items, while allocating less resource to foveal items (Bays & Husain, 2008). Therefore, simultaneous averaging of two features on spatially segregated sets could be more open to manipulations of attentional factors. Whereas, when feature pairs vary on same object sets as in our study, there could be an advantage of integration of two features on the same objects and viewers can accurately average those two features. In addition, our results are in parallel with Utochkin and Vostrikov's (2017) findings where they found similar precue and postcue errors for simultaneous averaging of size and numerosity. They presented these two features on the same object set and asked for immediate response after presentation of the displays. Viewers' errors on average size and numerosity estimations were similar across precue and postcue conditions (Utochkin & Vostrikov, 2017) This method could explain our parallel findings by suggesting that, when the two features for simultaneous averaging varied on the same set, attentional demands might be lower, and viewers can more effectively average two feature, compared to spatially separated sets.

Secondly, Emmanouil and Treisman (2008) presented their stimuli for 520 milliseconds, and there was 500 milliseconds of interval between stimuli and response in postcue conditions. This could cause the representation of the irrelevant feature in postcue conditions to decay over the time of display presentation and interval. Whereas, in mixed blocks of our experiment, there was no delay before the response, which might have helped summary representation of the both features to stay intact until response phase.

2.5 Experiment 3

Since Experiment 2 provided additional support for the feature-specific mechanisms of ensemble processing, in Experiment 3, we wanted to examine the temporal independence of size and orientation summarization processes. It is well known that in the visual processing hierarchy orientation feature is processed at early cortical stages, while size feature is processed at ventral stream (Whitney, Haberman & Sweeny, 2014). We expected participants to have higher averaging errors to increase as encoding duration decreased. We also expected the orientation feature to be averaged with less error than size feature at earlier durations, because of their different temporal position in visual processing hierarchy.

2.5.1 Method

2.5.1.1 Participants

Twenty-six undergraduate Bogazici University students participated in this experiment in return of course credits for their participation. We excluded data of one participant because of having orientation averaging errors more than 3 standard deviations away from the group mean, and another participant for having irrelevant responses. Therefore, we conducted the analyses on the data from 24 participants.

2.5.1.2 Material, stimuli and procedure

We used the same procedure and design as Experiment 2, except we manipulated so that displays were presented for either 50, 100, or 200 milliseconds. Encoding durations were intermixed throughout the trials of all blocks.

2.5.2 Results

For each participant, we calculated size and orientation averaging errors. We separately carried out a 2 x 3 repeated measures ANOVA analysis with single-mixed block condition and encoding duration as the variables, on size and orientation averaging errors. For the size errors, there was no main effect of single-mixed block conditions and encoding durations. Results showed that, there was no significant effect of single-mixed conditions on size averaging errors ($F(1,23) = 0.59$, $p = .810$, $\eta_p^2 = .003$), and errors were similar across different durations ($F(2,22) = 2.64$, $p =$

.093, η_p^2 = .194) (Figure 9). A similar pattern emerged for orientation averaging errors. There was also no difference between errors of single and mixed orientation averaging tasks (F(1,23) = 1.69, p = .206, η_p^2 = .069), and orientation averaging errors were similar across duration conditions (F(2,22) = .15, p = .866, η_p^2 = .013) (Figure 10). In size averaging task, the duration X single-mixed task interaction effect was not significant F(2,22) = 3.14, p = .063, η_p^2 = .222; while there was a significant interaction effect for orientation averaging task F(2,22) = 6.43, $p < .05$, η_p^2 = .369.

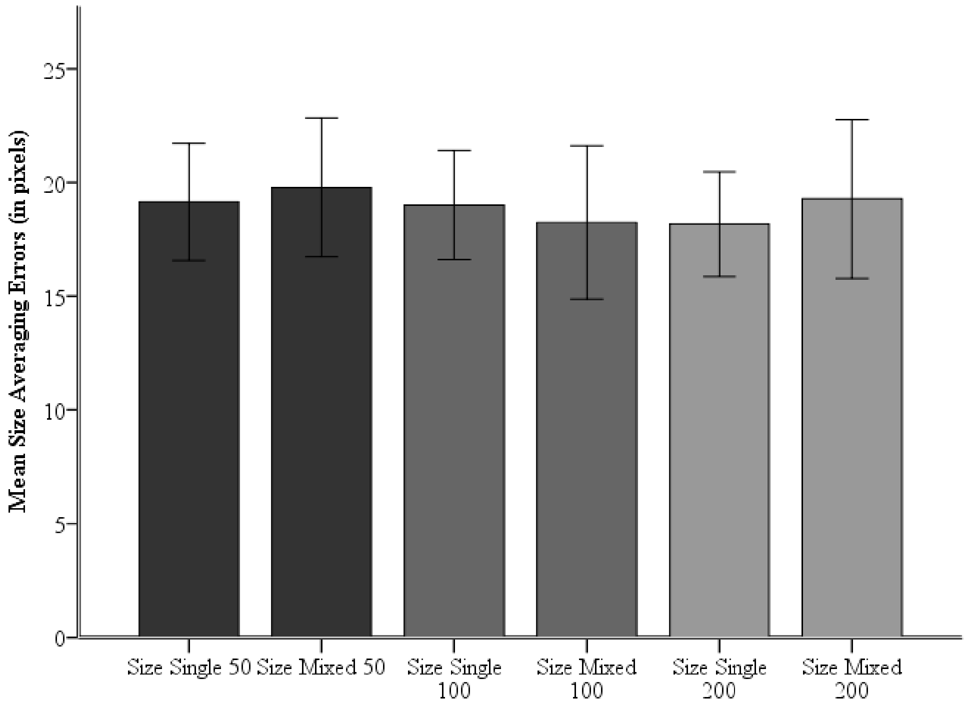

Figure 9. Mean size averaging errors across single and mixed conditions and encoding durations. Error bars indicate ±2 standard error of the mean.

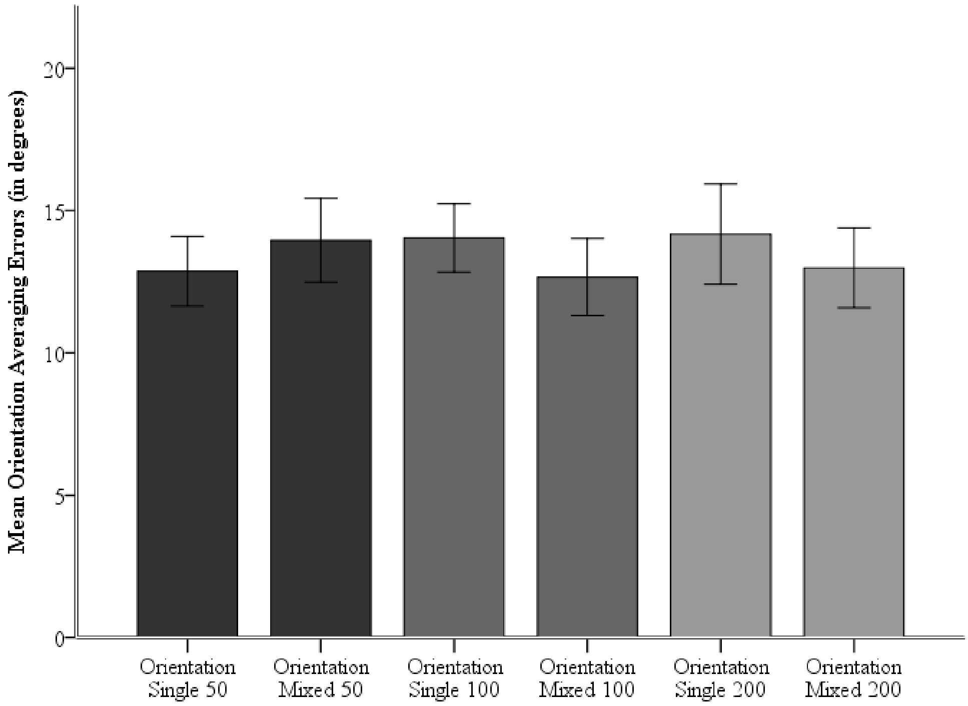

Figure 10. Mean orientation averaging errors across single and mixed conditions and encoding durations. Error bars indicate ±2 standard error of the mean.

We also analyzed the relation between participants' performance on size and orientation averaging tasks, across all single-mixed conditions and durations. There were no significant correlation between participants' errors on single block size and orientation averaging tasks across all encoding durations (all *r*s < .244, all *p*s > .251). In addition, there was no significant correlation between participants' mixed block size and orientation averaging errors across all encoding durations (all *r*s < |.133|, all *p*s > .594).

2.5.3 Discussion

Experiment 3 demonstrated that reducing encoding duration to 50 milliseconds did not have a significant impact on the error rates in neither of the size and orientation features. This was the case both for single or mixed block tasks. Participants made similar levels of error on size and orientation averaging tasks across different encoding (50, 100, and 200 milliseconds). In addition, there was no correlation averaging errors for size and orientation averaging tasks both for single and mixed averaging performances for all encoding durations. These results are consistent with results from earlier experiments, which argued that there are independent statistical summary mechanisms for size and orientation, and strengthened our findings about independent feature-specific statistical summary mechanisms for size and orientation features.

We expected increased errors at shorter encoding durations considering Jacoby, Kamke and Mattingley's (2013) findings on decreased influence of masked items on average orientation estimation; specifically we expected an early processing advantage for orientation averaging, since orientation feature is processed at early cortical stages of visual system. Jacoby, Kamke and Mattingley (2013) asked participants to extract average size or orientation of an object set. When they masked a subset of the object set, contribution of masked objects to average size and orientation estimations reduced. Their findings suggested late processing advantage for statistical summary representations. Even though we did not find increased errors at earlier encoding durations for both size and orientation features, this could not be enough to indicate possible early processing advantage for averaging mechanisms of those features, because of the experimental design differences between our and Jacoby and his colleagues' (2013) study. They used an even shorter encoding duration (30 milliseconds) and used a method with masking paradigm. Therefore, 50 milliseconds could be enough to extract an accurate statistical summary representation. This could indicate that even if size and orientation features do not have hierarchically different places in statistical summary processing mechanisms, there could be still complete independence in their averaging mechanisms.

CHAPTER 3

GENERAL DISCUSSION

This study investigated whether there are feature-specific statistical summary mechanisms in visual information processing. Across a series of experiments, we compared error rates in orientation and mean size estimation tasks and concluded that once attentional task demands are eliminated from designs, there was no longer a significant correlation between errors in these domains. These findings suggest that processing of mean size and orientation are independent and supportive of feature-specific summary mechanisms in vision. Our findings are in line with earlier work suggesting that there are different summarizing mechanisms operating at different levels (Haberman, Brady & Alvarez, 2015), and also previous findings from our lab suggesting that visual and spatial averaging mechanisms may be independent (Uner, Mutluturk & Boduroglu, 2014; Yildirim & Boduroglu, 2015).

While our findings do not speak to the neural mechanisms of the feature-specific summary mechanisms, they are nonetheless consistent with proposals on the neural basis summary mechanisms. For instance, Whitney, Haberman & Sweeny (2014) speculated that there could be multilevel statistical summary processors in different regions of the brain, specifically orientation feature could be averaged at early cortical stages and size feature could be averaged in ventral pathway of the visual system. In addition, Myczek and Simons (2008) argued that early visual system does not have specific receptors for size, unlike orientation feature. Therefore, our findings about separate feature-specific summary mechanisms for size and orientation features could be explained by the neural underpinnings of these processes. Future research with a dissociative experimental design and functional magnetic resonance imaging (fMRI) tools could shed some light on neural correlates of averaging processes of size and orientation features.

The findings on separate summary mechanisms, is conceptually meaningful given evidence for models that propose that visual short-term memory capacity is limited by capacity of independent feature stores. For instance, Bays, Wu and Husain (2011) showed that there is a strong independence between storage of color and orientation features in visual working memory. They gave viewers an array of objects with various orientations and colors. Viewers' task was to report color and orientation of the spatially indicated object by retrieving the representation from their visual working memory. They found that viewers' absolute error rates for estimated orientation and color features were uncorrelated. Their results suggested independent feature-specific storage and maintenance mechanisms for basic visual features in visual working memory. In a similar vein, Pasternak & Greenlee (2005) demonstrated that there is selective decay in visual working memory. Another working memory study where viewers were asked to recall color and location features showed that, distractors for each feature only interfered with the items that belong to the same dimension, suggesting that color and location features are separately processed (Pasternak & Greenlee, 2005). These studies suggest that there are independent feature-specific mechanisms for storage of basic features of visual information. Our study adds to that with independent feature-specific mechanisms for low-level visual features at the encoding stage of visual information processing where viewers represent object sets as statistical summary representations.

We found that viewers could efficiently process two low-level visual features as statistical summary representations with independent feature-specific mechanisms, yet there remains a question: Is there a limit to the number of features that we could accurately represent during statistical summary processing? As far as we know, studies on simultaneous statistical summary processing used mostly two-feature averaging settings, as the presented set of studies. This question could help us to understand these two issues more clearly: First, by knowing if there is a capacity limitation to the number of features for statistical summary processing, we could make further interpretations about unitary or multilevel mechanisms. If people could efficiently average more than two features in simultaneous settings, there should be higher capacity for averaging because of having multilevel independent statistical summary processors. On the other hand, if people could not efficiently average more than two visual features simultaneously, that could be a result of having a central statistical summary processor. Second issue is, whether capacity of statistical summary processing is limited with a maximum number of features that viewers could represent as summaries at a given time. Further studies should use designs with conjunction of 3, 4 or even 5 features in simultaneous statistical summary tasks. If viewers' performances decrease as the number of features increase it could be still possible to say that there might be a central statistical summary processor with a capacity to allocate resources only to a certain number of representations. On the other hand, if viewers' performances on these tasks stay stable as the number of features increase, we could more confidently say that there are multi-level independent statistical summary mechanisms.

REFERENCES

Albrecht, A. R., & Scholl, B. J. (2010). Perceptually averaging in a continuous visual world extracting statistical summary representations over time. *Psychological Science*, *21*(4), 560-567.

Albrecht, A. R., Scholl, B. J., & Chun, M. M. (2012). Perceptual averaging by eye and ear: Computing summary statistics from multimodal stimuli. *Attention, Perception, & Psychophysics*, *74*, 810–815.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122-131.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157-162.

Attarha, M., Moore, C. M., & Vecera, S. P. (2014). Summary statistics of size: Fixed processing capacity for multiple ensembles but unlimited processing capacity for single ensembles. *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), 1440.

Attarha, M., & Moore, C. M. (2015a). The capacity limitations of orientation summary statistics. *Attention, Perception, & Psychophysics*, *77*(4), 1116-1131.

Attarha, M., & Moore, C. M. (2015b). The perceptual processing capacity of summary statistics between and within feature dimensions. *Journal of Vision, 15*(4), 9-9.

Bauer, B. (2009). Does Stevens's power law for brightness extend to perceptual brightness averaging? *The Psychological Record*, *59*(2), 171.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science,* 321(5890), 851-854.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia,* 49(6), 1622-1631.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22(3), 384-392.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393-404.

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*(7), 891-900.

Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A, 18*(5), 1016-1026.

de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *The Quarterly Journal of Experimental Psychology*, *62*(9), 1716-1722.

Emmanouil, T. A., & Treisman, A. (2008). Dividing attention across feature dimensions in statistical processing of perceptual groups. *Perception & Psychophysics, 70*(6), 946-954.

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, *144*(2), 432.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751-R753.

Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience*, 4(2), 127.

Hubert-Wallander, B., & Boynton, G. M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of Vision*, *15*(4), 5-5.

Jacoby, O., Kamke, M. R., & Mattingley, J. B. (2013). Is the whole really more than the sum of its parts? Estimates of average size and orientation are susceptible to object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 233.

Maule, J., Witzel, C., & Franklin, A. (2014). Getting the gist of multiple hues: metric and categorical effects on ensemble perception of hue. *Journal of the Optical Society of America A*, *31*(4), A93-A102.

Morgan, M., Chubb, C., & Solomon, J. A. (2008). A 'dipper'function for texture discrimination based on orientation variance. *Journal of Vision*, *8*(11), 9-9.

Mutluturk, A. & Boduroglu, A. (2016). *Spatial summary statistics of a visual display.* Manuscript submitted for publication.

Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Attention, Perception, & Psychophysics*, *70*(5), 772-788.

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

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, 6(2), 97.

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., & Whitney, D. (2013). Humans use summary statistics to perceive auditory sequences. *Psychological Science, 24*(8), 1389-1397.

Robitaille, N., & Harris, I. M. (2011). When more is less: Extraction of summary statistics benefits from larger sets. *Journal of Vision, 11*(12), 18-18.

Semizer, Y. & Boduroglu, A. (2016) *Representing Visual Variance.* Manuscript submited for publication.

Uner, O., Mutluturk, A. & Boduroglu, A. (November, 2014). The impact of video game playing on the extraction of statistical summaries in visual displays.

Poster presentation at the *Poster presentation at the Annual Meeting of the Psychonomics Society,* Long Beach, CA.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*(3), 498-505.

Whitney D., Haberman J., Sweeny T. D. (2014). From textures to crowds: Multiple levels of summary statistical perception. Werner J. S., Chalupa L. M. (Eds.), *The New Visual Neurosciences*. Cambridge, MA: MIT Press.

Yildirim, A. & Boduroglu, A. (November, 2015). Binding features in spatial statistical summary representations. Poster presentation at the *Annual Meeting of Object, Perception and Attention (OPAM),* Chicago, IL.