

Adaptation to variance generalizes across visual domains

Article (Accepted Version)

Maule, John and Franklin, Anna (2020) Adaptation to variance generalizes across visual domains. *Journal of Experimental Psychology: General*, 149 (4). pp. 662-675. ISSN 0096-3445

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/85700/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Adaptation to variance generalises across visual domains

John Maule *

Sussex Colour Group, School of Psychology, University of Sussex

J.Maule@sussex.ac.uk

Pevensey II 5B6
School of Psychology
Pevensey 1 Building
North-South Road
Falmer
BN1 9QH

Anna Franklin

Sussex Colour Group, School of Psychology, University of Sussex

A.Franklin@sussex.ac.uk

Pevensey I 2B8
School of Psychology
Pevensey 1 Building
North-South Road
Falmer
BN1 9QH

Previous dissemination:

Data from this manuscript were presented in a poster at the 24th Symposium of the International Colour Vision Society (ICVS) in Erlangen, Germany. The poster is available on ResearchGate (doi: 10.13140/RG.2.2.33292.41606).

* Corresponding author

Word count: 9,741

Abstract

Processing the vast amount of visual information available from the world ought to pose a significant challenge to the brain. One of the ways in which the brain appears to encode the structure inherent in the world is through summary statistical representations (e.g. mean size, colour etc). This study investigates whether variance perception can be adapted for colour, and then whether the variance adaptation aftereffects generalise from colour to another visual domain. In a series of four experiments we find aftereffects reflecting adaptation to the variance of hues in an ensemble – such that prolonged exposure to a highly variable ensemble of hues makes subsequent ensembles appear less variable in hue. We also found that this effect partially generalised to the perception of orientation variance – adaptation to highly variable colour ensembles made subsequent ensembles also appear less variable in orientation. This is a novel demonstration of adaptation aftereffects reflecting processing of visual ensemble information across domains. The results could imply a neural mechanism encoding visual variance which is not selective to the domain from which the variance signal is derived. This mechanism may form the basis for cross-domain visual comparisons, and may play a role in predictive coding, enabling the brain to calibrate to the complexity of the visual environment.

Keywords: ensemble perception, predictive coding, adaptation, vision

Adaptation to variance generalises across visual domains

The visual environment is highly complex, and the human visual system processes a constant stream of information across various domains (e.g. colour, lightness, orientation, size, texture) to provide us with the sensation of a seamless and holistic experience of the surfaces and objects around us. Compressing visual information into summary statistics is one of the ways in which our visual system takes advantage of regularities in the structure of the environment (Dakin, 2012), and optimises visual processing (Alvarez, 2011).

Many experimental studies have shown that observers are able to average the visual features of a set of similar exemplars. This phenomenon is known as ensemble perception and has been demonstrated for many different visual domains, including colour (Chetverikov, Campana, & Kristjansson, 2017; Maule & Franklin, 2015, 2016; Maule, Stanworth, Pellicano, & Franklin, 2016; Maule, Witzel, & Franklin, 2014; J. Webster, Kay, & Webster, 2014), orientation (e.g., Dakin, 2001; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), size (Ariely, 2001, 2008; Corbett & Oriet, 2011; Marchant & de Fockert, 2009; Simons & Myczek, 2008; Utochkin & Tiurina, 2014), brightness (Bauer, 2009), motion direction (Watamaniuk & McKee, 1998), motion speed (Watamaniuk & Duchon, 1992), facial expression of emotion (Haberman & Whitney, 2009, 2010), facial identity (de Fockert & Wolfenstein, 2009) and gaze direction (Sweeny & Whitney, 2014). Ensemble representations have been shown to be a driver of adaptation aftereffects. Following adaptation to ensemble arrays with larger mean sizes, subsequent arrays appear to have a smaller average size (Corbett, Wurnitsch, Schwartz, & Whitney, 2012). Similarly, the mean attractiveness of a set of faces causes a shift in attractiveness judgments consistent with adaptation aftereffects (Ying, Burns, Lin, & Xu, 2019). The mean expression of a serially-presented set of faces can affect subsequent judgments of expression, even under very rapid presentation where each face is presented for only 23.5ms (Ying & Xu, 2017).

Despite the broad range of stimuli for which observers are able to extract averages, and therefore the multiple levels in the visual system at which averaging may be taking place (for a comprehensive review see Whitney & Yamanashi Leib, 2018), there are a number of striking and common features which occur across domains. First, ensemble perception is quick – accurate visual averaging is intact even when a set is displayed for very short durations – typically 100-500ms. Second, the accuracy of visual averaging is unaffected by the number of elements in a set (Ariely, 2008; Marchant, Simons, & de Fockert, 2013; Simons & Myczek, 2008). Third, visual averaging is drastically affected by the heterogeneity of a set – mean estimates are more accurate and more readily formed when elements are more perceptually similar to one another (Haberman & Whitney, 2010; Im & Halberda, 2013; Maule & Franklin, 2015; Utochkin & Tiurina, 2014). Fourth, ensemble perception appears to be extremely efficient, with observers capable of extracting summary statistics for sets which exceed the 3-4 item limit of visual working memory (Ariely, 2008; Sweeny & Whitney, 2014; Yamanashi Leib et al., 2014), even in children (Sweeny, Wurnitsch, Gopnik, & Whitney, 2014). Whether this implies that ensemble perception is based on a holistic process, or might be the result of subsampling is an area of lively debate (Ariely, 2008; Attarha, Moore, & Vecera, 2014; Chong, Joo, Emmanouil, & Treisman, 2008; Marchant et al., 2013; Maule & Franklin, 2016; Simons & Myczek, 2008; Utochkin & Tiurina, 2014). Eye-movements and retinal location may be key factors in how ensemble statistics are encoded and whether encoding seems to support holistic or subsampling mechanisms (see Wolfe, Kosovicheva, Leib, Wood, & Whitney, 2015 for discussion). Fifth, ensemble perception is obligatory – perception is biased by the mean of the visual features in a display even when the task does not require explicit average judgments (Allik, Toom, Raidvee, Averin, & Kreegipuu, 2014; Corbett & Melcher, 2014; Maule et al., 2014; Oriet & Brand, 2013; Parkes et al., 2001). Similarly, evidence of residual visual averaging has been found even in patients whose perception of individual items is impaired by

neurological damage, in simultagnosia (Demeyere, Rzeskiewicz, Humphreys, & Humphreys, 2008), prosopagnosia (Leib, Puri, et al., 2012), and visual neglect (Lanzoni, Melcher, Miceli, & Corbett, 2014; Leib, Landau, Baek, Chong, & Robertson, 2012; Pavlovskaya, Soroker, Bonne, & Hochstein, 2015).

At present it is not clear whether these commonalities are as a result of similar organisation throughout the visual system, or whether there could be a domain-general mechanism which extracts and represents the summary statistics of the visual environment. Haberman, Brady and Alvarez (2015) have shown that individual averaging performance is correlated across domains for low-level properties of colour and orientation, and across the high-level properties of facial expression and identity, while performance in low-level averaging is not correlated with high-level averaging. This may imply a common mechanism, with shared sources of noise, for ensemble perception for basic representations of visual features, but distinct mechanisms for higher-level stimuli. Similarly, Yang, Tokita and Ishiguchi (2018) found that accuracy of estimating the mean or variance from ensembles was correlated across domains – performance extracting summary statistics from ensembles varying in size was related to performance for ensembles varying in orientation. In contrast they found that the accuracy of mean estimation was unrelated to variance estimation, suggesting that different neural mechanisms code the first- and second-order statistical representations, but on a domain-general basis.

In contrast to the intensity of research and discussion on extraction of the mean, there has been a more limited exploration of how the spread (e.g. variance) of stimuli might affect ensemble coding, or whether variability itself is directly encoded and is part of an observer's impression of a set. Variance discrimination has been shown for ensembles of faces (Haberman, Lee, & Whitney, 2015), size (Lau & Brady, 2018; Solomon, Morgan, & Chubb, 2011) and orientation (Morgan, Chubb, & Solomon, 2008). Norman, Heywood and Kentridge

(2015) showed that observers can be adapted to orientation variance – they displayed two ensembles of Gabor patches with matched mean orientation but differing variances and asked observers to judge which appeared more variable. They found that after prolonged exposure to higher variance in one retinal location and lower in another, observers showed a classic adaptation aftereffect. Responses were biased such that the location which had previously been associated with low variance was perceived as being more variable than the location associated with high variance, even when ensembles were of equal variance. Furthermore, they found that the effects were robust even when the mean orientation was orthogonal between the adaptation and test phases. That is to say that adaptation to variance around vertical still had an effect on perception of variance of Gabor patches distributed around horizontal. This, the authors argued, was evidence for “direct encoding” of variance in the visual system. This suggests that, rather than variance perception emerging from a mechanism pooling across neurons coding the specific orientations present, there must be a neural substrate of orientation variance perception which operates independently of the neurons coding the specific orientations present in an ensemble. Put another way it is not the variance around a particular mean which is being adapted, but a mechanism specifically coding the variance. This independence from central tendency means that the coding of variance is used across the whole representational space for orientation, providing a “measure” of the variability in the stimulus which can be compared from one instance (horizontal central tendency) to another (vertical central tendency). Similar generalisation from one stimulus representation to another has been reported for judgments of “volatility”, indicating adaptation effects on how changeable stimulus numerosity is over time can affect quite different representations of the same information (Payzan-LeNestour, Balleine, Berrada, & Pearson, 2016).

If a direct neural encoding for variance for orientation is present perhaps there is also the same mechanism for other domains. Furthermore, given that variance is a substrate for

facilitating comparison across different regions of representational space of orientation (e.g. comparing vertical sets to horizontal sets), perhaps variance encoding might also function as a domain-general measure of the variation in the current environment. There is no obviously meaningful way to compare the mean size (for example) to the mean brightness of two sets. In contrast the variability of a feature is comparable across domains – one set which varies in size can be said to contain more or less variation than another set which varies in brightness. Could there be a mechanism that encodes variance independently not only of the central tendency, but also independently of the stimulus domain?

The present study had two main aims – firstly, to see whether variance adaptation aftereffects, similar to those reported by Norman et al. (2015) can be established for another domain – hue. Investigating whether the variance adaptation effect is also found in other stimulus domains is important in furthering our understanding of the common properties of the ensemble coding mechanisms in the visual system. Hue is of particular interest in ensemble coding as it has the unusual property of being circular (consistent change in a single “perceptual direction” results in an eventual return to the starting hue). It is also a qualitative dimension to which humans readily apply categorical labels to communicate their experience. Despite this there is good evidence that hue ensemble averaging can still take place, albeit with constraints to the range over which averaging is robust (Maule & Franklin, 2015). Encoding the variance may be relevant to some domains, but not others, and there may be ecological reasons for this relating to the presence and information content of that signal in the visual environment. If differences are found between domains it is reasonable to assume that there are ecological reasons related to the properties of the visual diet explaining why summary statistical encoding functions as it does in each case. In the case of hue, the variance across the surface of an object may be indicative of particular states or properties (e.g. changes in homogeneity of the hues on fruit related to ripening), while the variance in a set might inform choices we make about how

to interact with the set (e.g. judging whether to hedge your bets by foraging from a bush with the greater variability in fruit ripeness, or a more homogenous bush). The second aim of this study was to address the question of whether variability adaptation aftereffects can generalise across stimulus domains. Specifically we investigated whether adaptation to hue variability biases judgements of orientation variability. Seeking cross-domain generalisation of variance adaptation is a strong test of whether domain general ensemble coding mechanisms are likely to exist. If adaptation effects do generalise across domains it could be suggestive of a neural locus or pathway in which ensemble variability, independent of stimulus, is encoded. Such a mechanism would have wide implications for our understanding of visual processing and how the brain codes and copes with the volatility of the visual environment. We focused on colour and orientation because it has been suggested that these two domains could share ensemble representation mechanisms (Haberma, Brady, et al., 2015). Furthermore, orientation variance aftereffects are already established, and perceptual colour spaces allow careful control of the ensemble properties, in order to be confident that any effects are not artefacts of poor stimulus control.

We have addressed these questions across four experiments. Experiment 1 and 2 attempted to establish variance aftereffects for hue and whether variance adaptation is independent of central tendency (following Norman et al., 2015). Experiment 3 explored the possibility that spatial arrangement may affect variance encoding, and experiment 4 tested whether colour variance aftereffects would generalise to orientation.

Experiment 1

This experiment sought to establish whether there is a variability adaptation aftereffect for hue. The same mean hue was presented in the ensembles throughout the experiment. This

means observers saw a particular range of hues during the task, and the mean hue of the ensembles presented were the same across adaptation flashes and test trials.

Method

Participants

Observers were recruited from the student and staff community at the University of Sussex and were naïve to the purpose of the experiments. All reported normal or correct-to-normal visual acuity and were assessed as having normal colour vision using Ishihara plates (Ishihara, 1973) and the Lanthony test (Lanthony, 1998). They received payment of £8 per hour of their time (*pro rata*), or course credit. The research protocol was approved by the University of Sussex Sciences and Technology Cross-Schools Ethics Committee (ER/JJM29/11) and European Research Council Executive Agency Ethics Review Board (COLOURMIND: 772193). Thirty-one observers (five male) took part (mean age 21.6 years) in Experiment 1. Observers were assigned to a condition – either “high” ($n = 12$), “low” ($n = 11$), or “neutral” ($n = 8$). Data from two observers (one high, one neutral) were excluded due to faults during testing.

Stimuli

The experiment consisted of a continuous train of ensembles, arranged in pairs to the left and right of a fixation point. Each ensemble contained four different hues, each hue represented three times. These 12 elements were arranged in a 4-by-4 grid subtending 6.25 degrees of visual angle. The corner cells of the grid were always empty (see Figure 2). Elements were circular uniform patches of colour each subtending 1.57 degrees of visual angle. Their

position was randomly jittered within their cell and colours were randomly distributed among the elements on each slide. The centre of each ensemble was positioned 6.25 degrees from the central fixation point.

Stimulus colours were specified in CIE $L^*u^*v^*$ (1976) colour space, which approximates perceptual uniformity for the average observer. In this space hue (the qualitative experience of colour often referred to using colour names, e.g. “red”, “blue”) and hue difference can be specified in terms of azimuth (polar angle) in degrees. The distance from the origin of the space (the white point, usually the background) provides a dimension, which approximates chromatic intensity or colour saturation, called “Chroma”. The third dimension is lightness (L^*), and is defined relative to the white point, where increments correspond to values above 100 and decrements to values below 100. All colours were selected from a circle on an equiluminant plane in CIE $L^*u^*v^*$ (1976) space, where $L^* = 95$ and Chroma = 80. The background (and whitepoint for $L^*u^*v^*$ calculations) was the monitor white at half maximum luminance (xyY(1931): 0.333, 0.353, 31.10) – a uniform grey.

Hue variability was manipulated by altering the u^*v^* hue angle difference between element colours (ΔH) – low variability ensembles $\Delta H = 15^\circ$, medium variability $\Delta H = 30^\circ$ and high variability $\Delta H = 45^\circ$. The mean hue was consistent across slides and across both ensembles, and consistent across the adaptation and test phases for each observer. A mean hue was assigned at random to each participant, with any hue in a full hue circle equally likely to be chosen as the mean for that observer. As such one participant might be exposed to ensembles all with a greenish mean hue, while another might see ensembles which all have a reddish, or bluish mean hue. Hence any effects found would likely reflect general properties of hue variance adaptation, rather than being related to any particular set of hues.

Apparatus

A 22-inch Mitsubishi DiamondPlus 2070SB CRT monitor, set to a resolution of 1280 × 960 pixels, 24-bit colour resolution, and a refresh rate of 85 Hz was used. Accurate rendering of the stimulus colours was achieved via calibration of the display with a Photo Research PR655 SpectraScan spectroradiometer. The experiment was carried out on a PC running MATLAB 2014a, with Psychtoolbox 3.0 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). A chin rest was used to constrain viewing distance to 57cm and responses were given using the keyboard. The experiment took place in a darkened room with walls painted in matt black, the desk was covered in a matt black cloth and both the monitor surround and LEDs on the keyboard were masked with black felt.

Design

A classic visual adaptation paradigm was used, involving a long adaptation phase interleaved with test phases, and generally following the protocol of Norman et al. (2015) (see Figure 1). Each “slide” consisted of two ensembles arranged to the left and the right of a central fixation point. Ensembles were composed of a cluster of filled circles, of varying hue. The hue of each element was selected to produce an ensemble with given mean and inter-element perceptual distance (variability). The mean hue of every ensemble was the same throughout the experiment, in both adaptation and test phases. This mean was chosen randomly for the individual observer. Observers were assigned to one of three adaptation conditions – “low”, “high” and “neutral”. Observers were also pseudorandomly assigned an “adapted side” – either left or right. Adaptation slides always consisted of an ensemble of medium variability on the non-adapted side, and an ensemble containing the adaptation variability (low or high) on the adapted side. Hence, during adaptation there was a consistent asymmetry between the variance

in hue present on one side of the slide compared to the other. In the neutral condition both ensembles in the adaptation slides had medium variability, so there was no asymmetry in the variance displayed on each side.

During test trials the non-adapt side of the slide would contain a medium variability ensemble, while the adapt side would contain one of seven test variability levels, from $\Delta H = 7.5^\circ$ to $\Delta H = 52.5^\circ$, in steps of 7.5° (method of constant stimuli – see figure 2). For the neutral condition one side was arbitrarily chosen to be the “adapted” side for each observer for the purposes of presenting the test stimuli, although both sides were actually equally variable during adaptation. On each test trial the observer was required to respond according to which side appeared more variable.

Each test trial was followed by three top-up adaptation slides. Long adaptation phases of 30 slides occurred at the beginning of the experiment and after every 25 test trials (Figure 3). Each test variability was presented 25 times, yielding a total of 175 trials per observer.

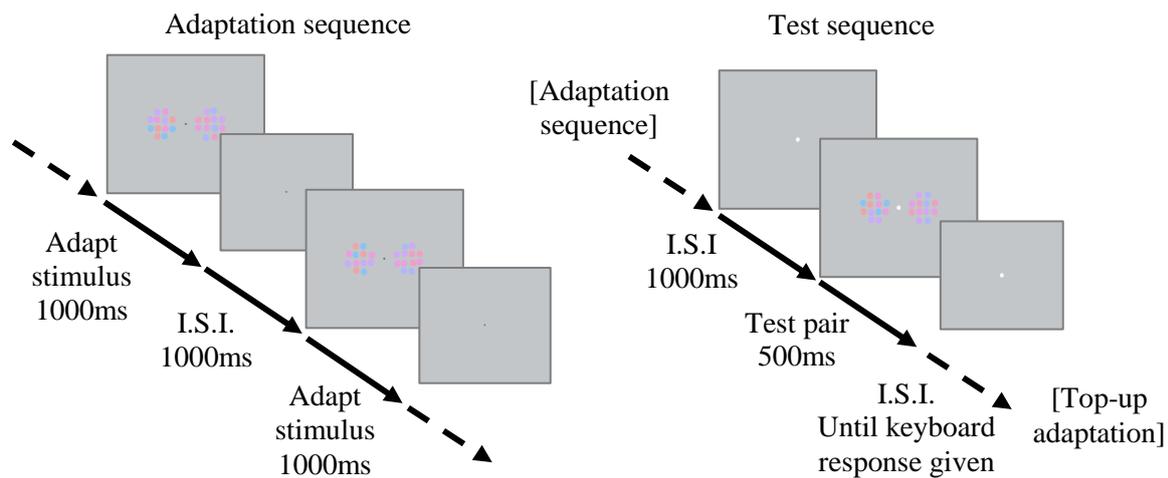


Figure 1: Example adaptation and test sequences (constant mean hue, as used for experiment 1). In this example there is high variability on the left side of the adapt stimulus. For each slide during an adaptation phase the ensemble elements are re-assigned their colours but the overall statistical properties (mean hue and variability level) is controlled. The fixation point was visible through the session, turning white to prepare the observer for a test trial. The general paradigm shown here was used for all experiments, although the spatial arrangement of stimuli shown here reflects that used for experiments 1 and 2 only. The colours presented indicate how slides, which all have the same mean hue, might appear. This is representative of the structure of experiment 1 only, whereas experiments 2-4 used a random mean hue for each slide (see details given in ‘Stimuli’ for each experiment).

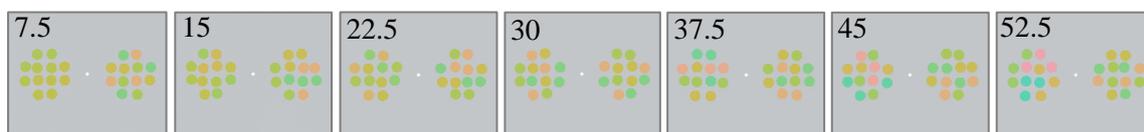


Figure 2: Example test stimuli illustrating variability levels for experiments 1 and 2. The variability “level” of the test ensemble is indicated by the numbers in the top left, and refer to the CIE $L^*u^*v^*$ hue angle difference between adjacent hues in the ensemble distribution. In these examples the test ensemble is on the left and is paired with an ensemble of the same mean hue and variability = 30. Note that for level 30 both ensembles contain the same hues and although their spatial arrangement is different, the overall (unadapted) percept is of equal variability on both sides.

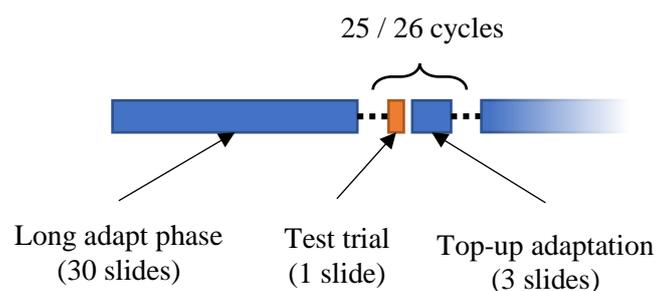


Figure 3: Schematic of experiment session structure. Each test trial was followed by a short phase of top-up adaptation, with longer adaptation phases after every 25 (experiment 1) or 26 (experiments 2-4) trials. These phases repeated until all the test trials were completed.

Procedure

Observers were briefed prior to the task and read written instructions on the screen prior to starting the experiment. The instructions emphasised that observers should keep their eyes fixed on the fixation point throughout. Practice trials using grey ensembles containing luminance variation were used to prepare the observer for the pacing and structure of the task, and also for the experimenter to assess whether the observer understood the task adequately. Visual feedback was given during practice trials, but never during the main experiment. The practice trials were relatively easy so if the observer responded correctly on fewer than 8 out of 10 trials the practice was re-run. The time spent reading the instructions and completing practice trials ensured adaptation to the white point.

Analysis

Psychometric functions were fitted to the data for each observer using a general linear model with a logit link function. The point of subjective equality (PSE – i.e. the level of variability at which the observer perceives the adapt side as equal in variability to the standard side) was defined as the 50% point on the psychometric curve. The predicted effect of adaptation is a shift in PSE towards the adapted stimulus. Therefore we expect PSE to be elevated in the high variability adaptation condition, relative to the low variability adaptation condition. If the conditions are separable in terms of PSE, and the shift is in the expected direction, this is evidence for an adaptation aftereffect. As we are predicting the direction of the change in PSE, all t-tests were one-tailed (following Norman et al., 2015). In addition we

used Bayesian statistics to assess the strength of the evidence for the alternative and null hypotheses (Dienes, 2014). We report log Bayes Factors (log BF) from Bayesian t-tests and ANOVA run using JASP (JASP Team, 2018) with the default Cauchy prior scaling of 0.707. By convention, log BFs below -1 indicate good evidence for the null hypothesis (no difference, or a difference in the opposite direction for one-tailed tests) and log BFs above +1 indicate good evidence for the alternative (Wetzels & Wagenmakers, 2012). Values between -1 and +1 are “insensitive”, although their sign still indicates the direction in which the evidence is pointing (negative towards the null, or positive towards the alternative).

Results & Discussion

One observer, from the neutral condition, was excluded due to showing extreme bias in their responses – pressing “right” on 80% of trials, resulting in a poor fit of the psychometric function to their responses. This resulted in final samples of $n = 11$ for the high condition, $n = 11$ for the low condition, and $n = 6$ for the neutral condition.

Estimates of point of subjective equality (PSE) were derived from the 50% point on the psychometric function for each observer. These were submitted to a one-way ANOVA with adaptation level (high, low, neutral) as the between-subjects factor. This revealed a significant main effect of adaptation level on PSE ($F(2,25) = 32.95, p < .001, \eta^2 = .725$). The log Bayes Factor was 11.89, indicating good evidence for a difference. Post-hoc t-tests examining the mean differences between each condition pairwise revealed that all three were significantly different from one another (smallest $t = 3.13$, largest $p_{\text{Bonferroni}} = .013$, smallest $d = 1.81$). Log Bayes Factors associated with these tests also indicated good evidence for a difference (smallest logBF = 2.58).

We observed that the slope of the psychometric functions also seemed to vary with adaptation condition (see Figure 4). Individual slopes were also submitted to a one-way ANOVA with adaptation level (high, low, neutral) as the between-subject factor. This revealed a significant main effect of adaptation level on slope ($F(2,25) = 6.17, p = .007, \eta^2 = .330$). The log Bayes Factor was 2.07, indicating good evidence for a difference. Post-hoc t-tests revealed that slopes in the high variability adaptation condition were significantly shallower than in the low condition ($t(20) = 2.81, p_{\text{Bonferroni}} = .029, d = 1.21, \log(\text{BF}) = 1.62$) and the neutral condition ($t(15) = 3.09, p_{\text{Bonferroni}} = .01, d = 1.49, \log(\text{BF}) = 1.62$). There was no significant difference between the mean slopes in the neutral and low conditions ($t(15) = 0.73, p_{\text{Bonferroni}} > .99, d = 0.39, \log(\text{BF}) = -0.64$).

We found adaptation aftereffects in the expected direction (Figure 4). Exposure to high variability ensembles caused subsequent ensembles to look relatively less variable, and exposure to low variability ensembles caused subsequent ensembles to look relatively more variable.

A post-hoc power analysis using GPower (Faul, Erdfelder, Lang, & Buchner, 2007) revealed that in order to detect an effect of the size found in this experiment a total sample of 15 (i.e. 5 per group) would be sufficient. Based on the difference between the high and low conditions alone, a total sample of 8 (4 in each condition) would suffice. Therefore, in the experiments that followed we recruited smaller groups and, from experiment 3 onwards, did not include the neutral condition.

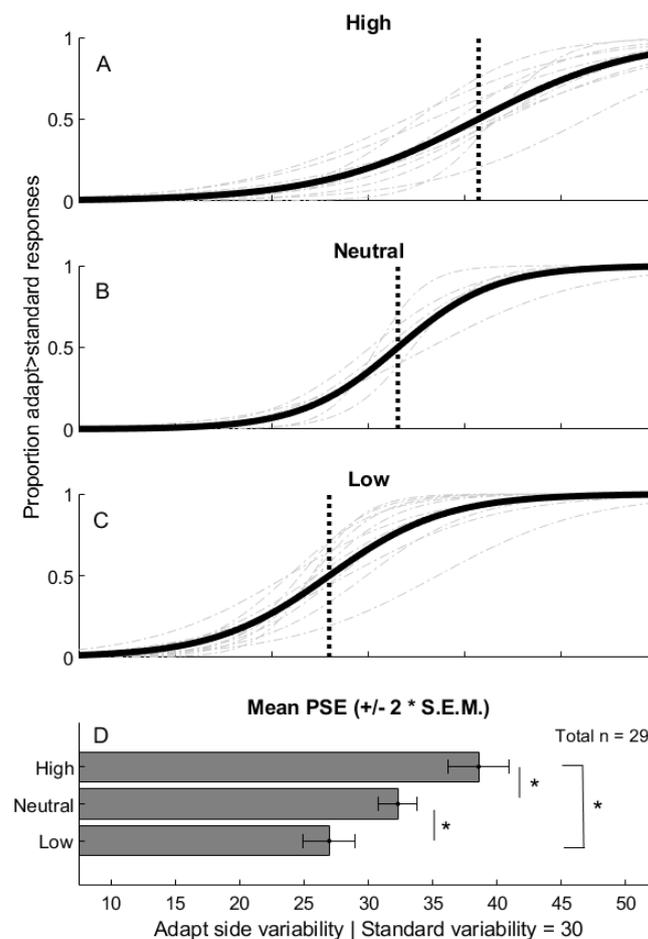


Figure 4: Results of experiment 1. On panels A to C dashed grey lines indicate individual psychometric functions while the solid line illustrates the mean psychometric function. Each panel displays the results from one of the three adaptation conditions. Vertical dashed lines indicate the position of the mean point of subjective equality (PSE). Panel D illustrates mean PSE for each condition. The ideal observer PSE (i.e. position of the PSE for an observer showing no bias and answering according to the objective variability level on each trial) is at 30 on the x-axis. Asterisk (*) indicates $p_{\text{Bonferroni}} < .05$ for independent t-tests between conditions linked by the bar.

Experiment 2

Experiment 1 established that hue variability judgments around a given hue can be biased by adaptation to variance around that hue. However, the striking finding from the Norman et al. (2015) experiments on orientation variance was that the effect generalised beyond the specific orientation adapted – such that adapting to variance around horizontal

would affect judgments of variance around vertical. Following adaptation to a given mean hue, as in experiment 1, aftereffects on variability judgments could occur as a result of changes in tuning or gain of neurons selective for the hues presented. As such, the aftereffects might not be the result of “direct” variance encoding, but may emerge instead from changes which reflect the consistent central tendency of the ensembles. To address this, in experiment 2 we allowed the hue presented on each slide to vary at random, such that the observer would see a full hue circle during the course of the experiment. Therefore any adaptation to the variability could not be the result of changes to hue-selective neurons, but would instead imply direct encoding of the variability without selectivity for the central tendency of the ensembles viewed.

Method

Participants

Twenty-two observers (four male) took part (mean age 22.2 years). Observers were assigned to a condition – either “high” ($n = 9$), “low” ($n = 7$), or “neutral” ($n = 6$). None had taken part in experiment 1.

Stimuli

Stimuli were as in experiment 1, except the mean hue was selected pseudo-randomly for each slide – i.e. both ensembles on the left and the right were of the same mean hue, but this varied from slide to slide. Therefore, across trials the observers were likely to see colours from around the whole hue circle, rather than just a small section of the hue circle as in experiment 1.

Design

The design was as in experiment 1, except two test variability levels were added to provide denser sampling around the neutral point. Nine test points were used, corresponding to variability levels of $\Delta H = 7.50^\circ, 15.00^\circ, 22.50^\circ, 26.25^\circ, 30.00^\circ, 33.75^\circ, 37.50^\circ, 45.00^\circ, 52.50^\circ$. Each test variability was presented 20 times, yielding a total of 180 trials per observer. Long adaptation phases were shown every 26 trials.

Results & Discussion

One observer, from the high condition, was excluded due to showing extreme bias in their responses – pressing “left” on 85% of trials, resulting in a poor fit of the psychometric function.

Estimates of point of subjective equality (PSE) were derived from the 50% point on the psychometric function for each observer. These were submitted to a one-way ANOVA with adaptation level (high, low, neutral) as the between-subject factor. This revealed no significant main effect of adaptation level on PSE ($F(2,19) = 0.05, p = .948, \eta^2 = .006$). The log Bayes Factor was -1.33, indicating good evidence for a lack of difference between the PSEs across adaptation conditions. As well as the absence of an adaptation effect, the slope of the psychometric functions was much shallower in experiment 2 ($M_{\text{all}} = -4.45, SD_{\text{all}} = 1.55$) compared to experiment 1 ($M_{\text{all}} = -2.68, SD_{\text{all}} = 1.02$) (Figure 5). Within the experiment, however there were no differences in slope across the three conditions ($F(2,19) = 0.36, p = .699, \eta^2 = .037, \log(\text{BF}) = -1.13$).

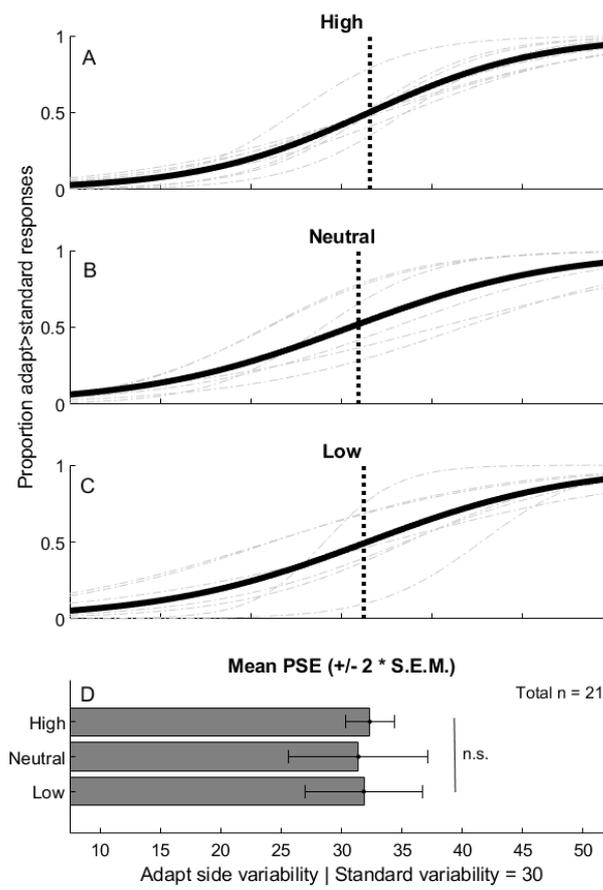


Figure 5: Results of experiment 2. Panels A-C illustrate individual and mean psychometric functions for each condition. Details are as given for figure 4, except (panel D) n.s. indicates $p > .05$ for ANOVA comparing mean PSE across the three conditions.

We found no evidence for adaptation aftereffects under the changing hue conditions in this experiment. The difference observed in the slopes of the psychometric functions suggest that observers' sensitivity was lower in experiment 2 than experiment 1. This could reflect the fact that in experiment 2 there was additional variability in the stimuli, as the mean hue changed with each slide. It is possible that direct encoding of variability does not apply to hue as it does for orientation, in which case the effects observed in experiment 1 may be the result of a kind of second-order variability estimation derived from the encoding of the hues present. If it is true that hue variability is not encoded independently of the mean this would present a contrast

to not just results from orientation ensembles but also recent findings of independent coding of the mean and variance for ensembles of faces which vary in attractiveness (Luo & Zhou, 2018). An alternative explanation for the lack of adaptation aftereffects is that the spatial arrangement of stimuli affects how summary statistics are extracted, which we addressed in experiment 3.

Experiment 3

Norman et al. (2015) used a relatively dense array of small Gabor patches, whereas experiments 1 and 2 used a few, relatively large elements as this size and density have previously been shown to result in reliable ensemble coding for hue (Maule & Franklin, 2015). It has been suggested that ensemble coding is closely related to texture perception (e.g. Dakin, 2001; Dakin & Watt, 1997), and as such ensembles with large elements which are easily individuated may not recruit summary statistical mechanisms as effectively as ensembles where the elements are small and densely packed, forming a percept which is more like a texture. Indeed, previous studies which have investigated summary statistical perception for colour have frequently used, with success, mosaics of tiny elements (e.g. Kimura, 2018; Kuriki, 2004; Li & Lennie, 1997; Sunaga & Yamashita, 2007). This experiment used a far more dense packing of elements in each ensemble, in order to address the question of whether dense texture-like arrays can drive variability adaptation more effectively than sparse arrays. The increased number of elements made it possible to specify the ensemble statistics in terms of the mean and standard deviation, with hues drawn from a Gaussian distribution.

Method

Participants

Fourteen observers (three male) took part (mean age 25.4 years). Observers were assigned to a condition – either “high” (n = 8) or “low” (n = 6). None had taken part in experiments 1 or 2.

Stimuli

Each ensemble contained 625 elements, arranged in a 25x25 grid subtending 4.70 degrees of visual angle. Elements were circular uniform patches of colour each subtending 0.19 degrees of visual angle. The hue of each element was selected from a Gaussian distribution, centred on a given mean hue and with a given standard deviation, the resulting hue angle was rounded to the nearest whole degree. The centre of each ensemble was positioned 4.70 degrees of visual angle from the central fixation point (see Figure 6).

Hue variability was manipulated by altering the SD of the normal distribution from which each element hue was selected. This SD is expressed in terms of $L^*u^*v^*$ hue angle in degrees. Low variability ensembles had $SD = 20^\circ$, medium variability $SD = 35^\circ$ and high variability $SD = 50^\circ$. These SDs were selected to approximate the SD of the ensembles in experiments 1 and 2 (low = 17.5° , high = 52.5°).

As in experiment 2, the mean hue was randomised from slide to slide, such that it was the same across both ensembles, but over the course of the experiment would result in a uniform sampling of the whole hue circle.

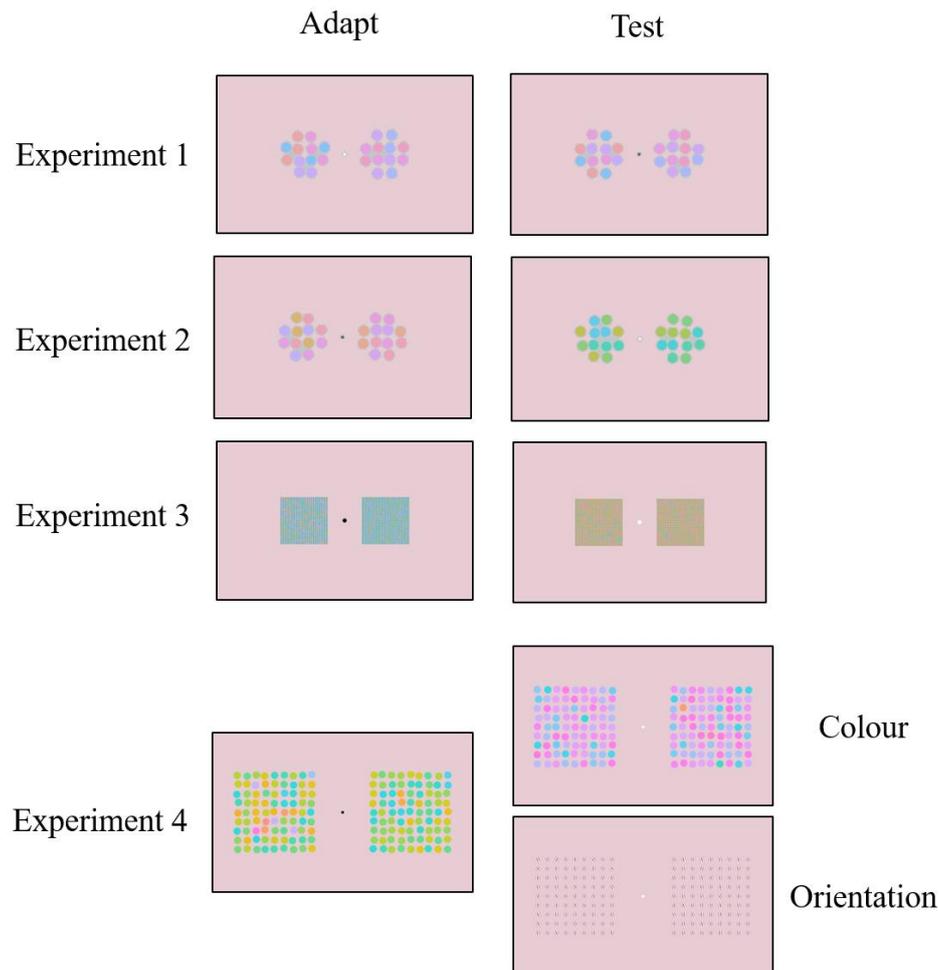


Figure 6: Spatial arrangement of stimuli for all experiments. In experiments 2-4 the adapt sequence involved the ensemble appearing with a random mean hue on each slide (but the same for the left and right ensembles), whereas in experiment 1 the same mean hue was presented on every slide. In experiment 4 test trials included either colour or orientation ensemble pairs. The task for all test trials was to judge which side appeared more variable.

Design

The design was the same as in experiment 2, except the test variability levels were specified in terms of the CIE $L^*u^*v^*$ hue SD: 20.0°, 25.0°, 30.0°, 32.5°, 35.0°, 37.5°, 40.0°, 45.0°, 50.0°.

Results & Discussion

One observer in the high condition was excluded due to showing extreme bias in their responses – pressing “left” on 86% of trials, resulting in a poor fit of the psychometric function. One observer in the low condition was excluded as their psychometric function had an extremely shallow slope, more than 3 SDs below the group mean slope.

Estimates of point of subjective equality (PSE) were derived from the 50% point on the psychometric function for each observer. An independent t-test showed that PSEs in the high adaptation condition were significantly greater than in the low adaptation condition ($t(10) = 7.67, p < .001, d = 4.49$) and the Bayes Factor indicated good evidence in support of this difference ($\log(\text{BF}) = 6.64$) (Figure 7). An independent t-test on the slope of the psychometric functions found that slopes were significantly shallower in the high condition compared to the low ($t(10) = 3.04, p_{\text{two-tailed}} = .013, d = 1.78$) and the Bayes Factor indicated good evidence in support of this difference ($\log(\text{BF}) = 1.49$).

Experiment 3 revealed a strong aftereffect of variability adaptation to ensembles consisting of many small elements of differing hues, forming a texture-like percept. This is in contrast to experiment 2, in which the ensembles had fewer but larger elements. In other regards the two experiments are very similar, and so it seems likely that there is an interaction between the spatial properties of an ensemble and the mechanisms used to encode its summary statistics. Crucially, this experiment also used a random mean hue for each slide, and so the variability aftereffect cannot be emerging from encoding the particular hues that are displayed, but rather must be “direct” encoding of the variability, as described for orientation by Norman et al. (2015). Again the data showed a reduction in sensitivity (slope) associated with exposure to higher variance in the ensembles.

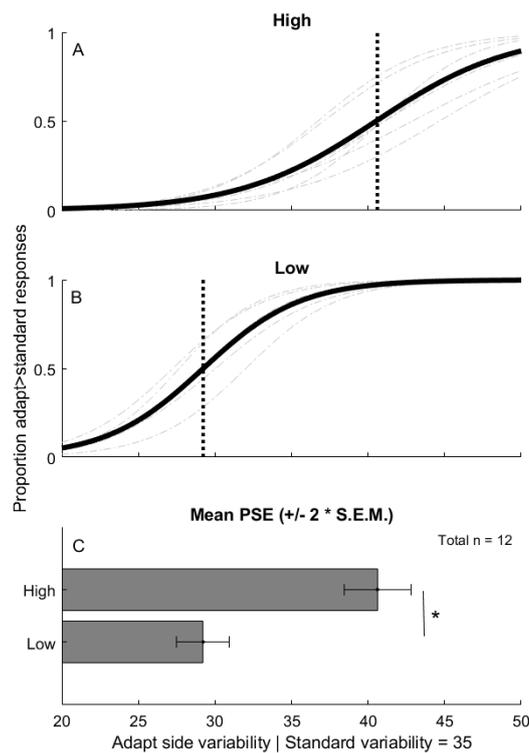


Figure 7: Results of experiment 3. Panels A and B show individual and mean psychometric functions. Other details as given for figures 4 and 5, except the ideal observer PSE is 35 for this experiment. Bar and asterisk (*) indicate $p_{\text{one-tailed}} < .05$ for the independent t-test comparing the high and low conditions.

Experiment 4

Experiment 1 and 3 have shown that observers adapt to variability in hue, and shown that this effect generalises across colour space. These experiments, and others in the literature summarised in the introduction, suggest that perception of summary statistics operates in a similar way across visual stimulus domains. Our next experiment investigated the as-yet unexplored question of whether there is direct cross-talk between domains in ensemble statistical judgments. This is important in the understanding of the neurophysiological basis for visual summary statistical processing. Cross-domain effects may be interpreted as evidence for a common neural substrate for visual processing of variability.

This experiment asked whether variability adaptation would generalise across distinct stimulus domains – does adaptation to variability in hue also affect judgments of orientation variability? The stimuli were similar to experiment 3, except test trials would present either an ensemble of hues, or an ensemble of tilted Gabor patches. The spatial features (number of elements and size) were equated across the two domains.

Method

Participants

Fifteen observers (six male) took part (mean age 20.1 years). Observers were assigned to a condition – either “high” ($n = 9$) or “low” ($n = 6$). None had taken part in experiments 1, 2 or 3.

Stimuli

Each ensemble contained 81 elements arranged in a 9-by-9 grid subtending 7.8 degrees of visual angle (see Figure 6). Elements subtended 0.6 degrees of visual angle. Their position was randomly jittered within their cell and colours were randomly distributed among the elements on each slide. The centre of each ensemble was positioned 7.8 degrees from the central fixation point. Since the effect of spatial arrangement on orientation variability judgments is unknown the size and number of elements in ensembles matched that used in Norman et al. (2015). This ensures that the orientation ensembles are appropriate but maintains a texture-like appearance for the colour ensembles, despite being a coarser array than in experiment 3.

Adaptation slides were as in experiment 3 – the elements were circles of colour, the hue of each element was derived from a normal distribution with a random mean, and a given SD. Therefore, the mean hue varied at random from slide to slide, although the mean hue of the two ensembles displayed at once on a single slide was always matched. However, in this experiment some test trials displayed ensembles varying in orientation, and some displayed ensembles of colours (see Figure 6). In the orientation ensembles the elements were Gabor patches with a spatial frequency of 4.2 cycles per degree and Michelson contrast of 90%. The tilt of each Gabor element was drawn from a normal distribution with a random mean and given SD. As in experiment 3, adaptation and test ensembles were paired with a “medium” variance ensemble ($SD = 35^\circ L^*u^*v^*$ hue angle), which is therefore the point of *objective* equality in the stimulus range, against which adaptation effects can be compared.

Design

Observers were assigned to either a high or low adaptation condition, defined as in experiment 3. Test variability levels for colour were 25.0° , 30.0° , 32.5° , 35.0° , 37.5° , 40.0° , $45.0^\circ L^*u^*v^*$ hue angle. For orientation trials the variability was scaled down to reflect the relatively finer discrimination of tilt angle compared to hue angle (e.g., M. A. Webster, De Valois, & Switkes, 1990). Test variability levels for orientation trials were $SD = 2^\circ$, 5° , 8° , 10° , 12° , 15° , 18° tilt from the mean orientation. On colour trials (adapt and test) the non-adapted side displayed an ensemble with $SD = 35^\circ L^*u^*v^*$ hue angle). On orientation trials the non-adapted side displayed $SD = 10^\circ$ tilt. Each test variability (7 colour and 7 orientation) was presented 15 times, yielding a total of 210 trials per observer. Colour and orientation test trials were pseudo-randomly interleaved.

As such, observers were adapted to differential variability of hue across locations, and tested on their perception of variability for both hue and orientation.

Results & Discussion

Two observers from the high condition were excluded as they showed extreme bias in their responses across the experiment – both choosing right on over 80% of trials. One observer from the low condition was excluded as their data showed a very poor fit for the psychometric function for the orientation condition ($r^2 = .01$).

The colour and Gabor conditions were analysed independently for differences. An independent t-test showed that PSEs for colour variability in the high adaptation condition were significantly greater than in the low adaptation condition ($t(10) = 2.89$, $p_{\text{one-tailed}} = .008$, $d = 1.67$), with good evidence in support of the effect ($\log(\text{BF}) = 2.00$). PSEs for orientation variability in the high adaptation condition were also significantly greater than in the low adaptation condition ($t(10) = 1.99$, $p_{\text{one-tailed}} = .038$, $d = 1.15$, $\log(\text{BF}) = 0.97$). It should be noted that the Bayes Factor indicates that the data support the experimental hypothesis that PSEs are greater in the high condition than in the low condition, but the strength of the evidence falls just short of the threshold of 1.0 to indicate the data is “sensitive” to a possible difference (Wetzels & Wagenmakers, 2012). We also analysed the slopes for each condition, finding that for psychometric functions of colour variance the slopes were significantly shallower in the high variance adaptation condition ($t(10) = 2.26$, $p_{\text{two-tailed}} = .047$, $d = 1.30$), with moderate evidence in support of the effect ($\log(\text{BF}) = 0.63$). For orientation there was no difference in slope between the high and low variance adaptation conditions ($t(10) = 0.36$, $p_{\text{two-tailed}} = .725$, $d = 0.21$, $\log(\text{BF}) = -0.72$).

One-sample t-tests were used to assess whether adaptation effects were observed in each condition individually. In the high condition the mean PSE for colour (40.71, SD = 5.21) was significantly greater than the “medium” variability level of 35 ($t(5) = 2.68$, $p_{\text{one-tailed}} = .022$, $d = 1.10$, $\log(\text{BF}) = 1.49$), and the mean PSE for orientation (12.82, SD = 2.74) was significantly greater than the “medium” variability level of 10 ($t(5) = 2.52$, $p_{\text{one-tailed}} = .027$, $d = 1.03$), with good evidence for the effect ($\log(\text{BF}) = 1.34$). In the low condition the mean PSE for colour (33.97, SD = 2.34) was not significantly less than the “medium” variability level of 35 ($t(5) = 1.08$, $p_{\text{one-tailed}} = .165$, $d = -0.44$), although the Bayes Factor suggested the data were insensitive ($\log(\text{BF}) = -0.06$). The mean PSE for orientation (10.44, SD = 1.03) was not significantly less than the “neutral” variability level of 10 ($t(5) = 1.06$, $p_{\text{one-tailed}} = .831$, $d = 0.43$), and the Bayes Factor indicated good evidence for no difference ($\log(\text{BF}) = -1.53$) (Figure 8).

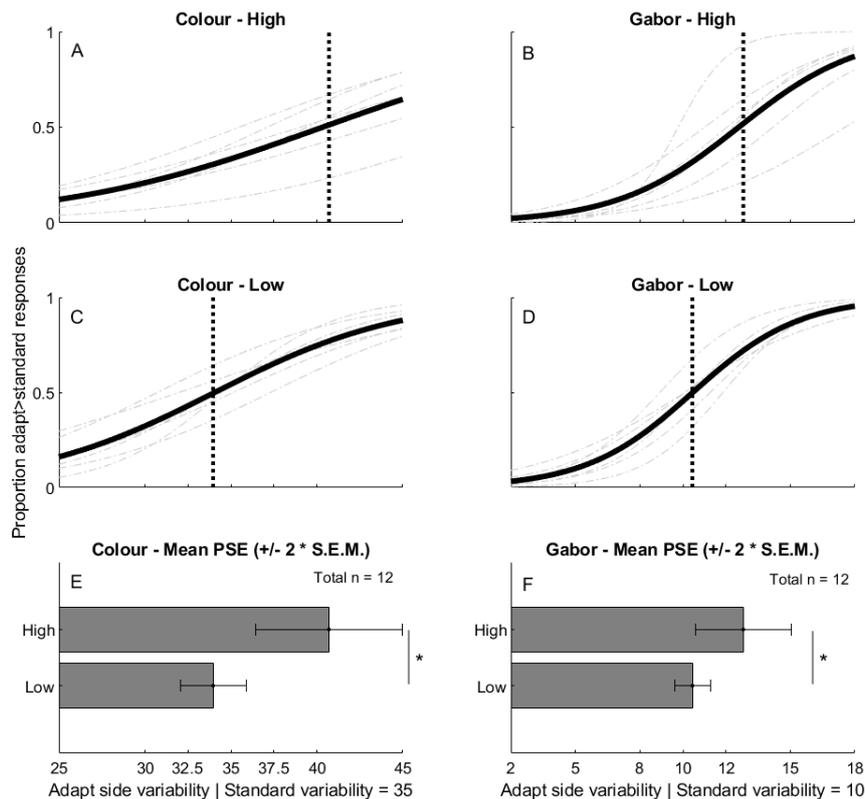


Figure 8: Results of experiment 4. Panels A-D show individual and mean psychometric functions for each of the four conditions (high-low across participants and colour-gabor within participants). Panels E and F indicate mean PSE differences (* indicates $p_{\text{one-tailed}} < .05$ for independent t-test comparing high and low conditions). Other details as given for figure 7, except for the Gabor (orientation) trials where the ideal (unbiased) observer PSE is 10.

The data show a novel aftereffect of adaptation to colour variance on apparent orientation variability – after adapting to high variance *colour* ensembles, perception of variability in ensembles of different *orientations* was reduced. These results suggest that there is generalisation of variability aftereffects from colour to orientation. This result may be interpreted as reflecting a domain-general encoding of variability. Such an interpretation would suggest that there is a neural locus encoding visual variability which is not selective to the stimulus type, but encodes how varied the environment is generally, pooling signals from different visual domains. The significance of this is further addressed in the general discussion.

The data were not unequivocal in their support of this idea, however. Aftereffects were not found in the low variance adaptation condition, in either colour or orientation variance judgements. In experiments 1 and 3 aftereffects were found for both low and high conditions, and so it is not immediately obvious why this pattern of results was not observed in the low variance condition in this experiment. Moreover it was not only the orientation aftereffect which was absent in the low condition, but also the effect on hue variability judgments. Previously, in experiment 3, the effect of adapting to low variance was equal in magnitude to adapting to high variance. Although there is a difference in spatial configuration, from the dense textures of experiment 3, with many small elements, to the ensembles in experiment 4 with larger elements (see Figure 6), the only other difference is the introduction of test trials with orientation variability, differences which we would not expect to cause the low variance adaptation aftereffect to be selectively diminished. We did observe a shallower slope of the psychometric functions for colour in experiment 4 compared to 3, although they were actually shallowest for the high condition. This difference suggests that observers had a greater range of uncertainty, or perhaps found the task harder when the test stimulus was varying between hues and Gabor patches. We speculate that this greater uncertainty may affect summary statistical processing differently for low and high variance adaptation. Under the more demanding task in experiment 4, in which the observer's task switches between judging hue variability and orientation variability – it may be that greater variability is more salient, driving summary representation more strongly and so variability encoding is adapted more readily.

Another possibility is that observers are biased in both conditions towards the adapting stimulus – a trend that can also be seen in experiment 1. This would have the effect of elevating the PSE in both conditions – making the high PSE an overestimate of the adaptation effect and the low PSE an underestimate. The cause of such a bias may be the way in which the paradigm has been designed – during the experiment the “non-adapted” side remains at a fixed variability

level throughout, while the adapted side alternates between a fixed variability during adaptation phases and various different levels of variability during test phases. Therefore across time there is more *change* in variability on the adapted side, in both conditions. When the task is more difficult observers may confuse this variability signal for the signal coming from the stimulus currently being shown. This would result in an overall elevation in the number of responses indicating that the observer perceives the adapt side as more variable, and so would bias the PSE across both conditions. Since all observers, regardless of condition, responded according to which ensemble appeared *more* variable any such bias would affect all conditions and does not confound the interpretation of the results. This possibility does not affect the fact that the PSEs of the low and high adaptation groups were statistically different – indicating that the adaptation procedure was affecting the groups in different ways.

General Discussion

This study presents evidence that adaptation to the variance for one visual feature can affect variability judgments of another visual feature, under certain conditions. One interpretation of this result, which would be of particular theoretical significance, is that there is a common mechanism encoding visual variability, and it is the activity of the neurons comprising this all-purpose variability measurement mechanism which are being modulated by the adapting stimulus to induce the aftereffect. This study also presents novel evidence for the encoding of variability in hue, independently of the central tendency (i.e. when mean hue varies on every slide – experiment 3 and 4). These findings support the notion that the brain's extraction of summary statistics of colour goes beyond the central tendency (e.g. mean) to also include representation of the spread of hue. This suggests that the visual system directly

encodes variability of hue, as it does orientation (Norman et al., 2015), for each ensemble presented.

If the visual system is to make use of sensory information from one domain to tune perception of another domain then variability is a good candidate. The central tendency of a distribution is only meaningful within its given domain – there isn't a meaningful comparison between an orientation ensemble with a mean of 15° and a colour ensemble with a mean hue which is reddish-orange. In contrast, variability is more comparable – particularly so for domains such as hue and orientation, where the differences are qualitative and finite, meaning that variability can be compared relative to the maximum variability possible (i.e. full hue circle or gratings from 0° to 180° tilt). In this way, variability could be a common currency to compare the information of the signals coming from different visual domains. However, why the brain should operate in this way is an open question. As mentioned, one possibility is that the carry-over of variability aftereffects from colour to orientation is the result of a shared neural substrate for coding variability in ensembles. If there is a single seat of ensemble coding across all visual domains we should expect variability aftereffects of colour to carry-over to domains as diverse as size and facial expression. The connection between accuracy in performing particular visual statistical judgments (mean or variance) across domains has precedence in the literature (Haberman, Brady, et al., 2015; Yang et al., 2018, but see also Anobile et al., 2018). This study contributes further evidence for cross-domain interactions and connections, and further support for a domain-general variance-encoding mechanism. Further research following the approach of experiment 4 and exploring inter-domain aftereffects across ensembles modulated in high and low level stimuli would help to further explore this question about the domain generality of summary statistical neural mechanisms. Studying the pattern of association and dissociation between different summary statistics and different domains will be crucial to this programme of research.

Another possible account for the partial carry-over of variance aftereffects from colour to orientation is that this effect is the result of the adaptation in the activity of populations of neurons which code for the conjunction of colour and orientation. Such neurons are well-established in the literature (e.g., Moran & Desimone, 1985; Zhaoping & Zhe, 2012) and are tuned to respond to a particular hue in conjunction with a particular orientation. There is also psychophysical evidence for the conjunction of these domains in processing (e.g., Holcombe & Cavanagh, 2001; McCollough, 1965). Therefore, there is a known neural substrate for encoding colour-orientation conjunctions, which may be the source of the variance signal which drives an adaptation aftereffect in orientation when the adapting stimulus is colour. This domain-specific (the “domain” here being colour-orientation conjunctions) mechanism may have the advantage of parsimony over postulating a domain-general mechanism which encodes variance from ensembles of any domain. However, the results of experiment 4 cannot be simply explained by the activity of colour-orientation tuned neurons, for the same reason that the hue variance aftereffects in experiment 3, and the orientation aftereffects found by Norman et al. (2015) cannot be explained by the simple encoding of those features. These results entail a “direct” encoding of variance to account for the adaptation aftereffects observed. It is only with further investigation into the extent and limits of within- and cross-domain variance adaptation aftereffects, that we may begin to understand whether they are related to the low-level encoding of specific domain conjunctions or could be indicative of domain general variance processing. Future research may also consider what purpose cross-domain comparisons might serve. Encoding variability could play an important role in calibrating the brain to the volatility of the environment (de Gardelle & Summerfield, 2011; Michael, de Gardelle, & Summerfield, 2014), or supporting metacognitive confidence in perceptual judgments (Suarez-Pinilla, Seth, & Roseboom, 2018). However if the brain is to do this optimally there should only be cross-talk between domains if the variability in one domain is

predictive of the variability in another. Therefore, finding cross-talk between any two stimulus domains raises a theoretically important question – is this a functional adaptation to correlations that are present in the visual environment, a by-product of incomplete separation of the signals derived from the visual stream or the result of a shared neural substrate which detects and encodes salient summary statistics across domains? We know that the brain is sensitive to correlations between some visual domains from contingent aftereffects – a striking and enduring example is the McCollough Effect (McCollough, 1965), in which prolonged exposure to a coloured grating (e.g. red and black horizontal) can induce an illusory colour in achromatic gratings, selective to the orientation of the grating (e.g. black and white horizontal grating appears greenish). To build up a McCollough aftereffect one must be exposed to the consistent contingency between orientation and colour. In contrast, in the present study the adapting stimulus doesn't include any contingent orientation signal, and yet aftereffects were found to carry over to that domain. An influence of colour variability on orientation variability perception implies that these image statistics are correlated in the natural world – scenes which vary more in hue should also vary more in orientation. This is a substantive claim which requires further research to address properly. The general principle, however, is that cross-domain effects could help demonstrate which visual features are correlated in the natural world, and therefore this paradigm could be used to guide investigations into the tuning of the neural encoding of visual scenes and the roles that cross-domain correlation plays in the processing of natural image statistics.

It is noted that the variability adaptation aftereffect shown in experiments 1, 3 and 4, where exposure to highly variable colour sets biases the colour variability perceived in subsequent sets, is reminiscent of the effects of contrast adaptation. When exposed to high contrast (i.e. highly saturated, intense) colours subsequent mid-contrast colours will appear lower in contrast (i.e. less intense, more similar to grey). Contrast adaptation is selective to the

direction of contrast modulation present in the adapting stimulus in terms of the cone-opponent mechanisms (e.g. M. A. Webster & Mollon, 1994). Importantly, however the effect demonstrated in the present study cannot be explained as a form of contrast adaptation. In experiments 3 and 4 the mean hue of the adaptation ensembles is assigned at random. That means that during the course of the experiment the observer is presented with colours from around the full hue circle, on both sides of the fixation point. Over the course of the experiment the net effect is equal stimulation of the cone-opponent mechanisms – i.e. there is no difference in the amount or polarity of low-level colour contrast presented as a result of the manipulation of variability. Furthermore, colour contrast adaptation could not account for the effect of adaptation to hue variability on judgments of orientation variability.

Another important consideration when interpreting this study, and indeed any study of adaptation aftereffects, is the possibility for the results to reflect only decision bias, rather than neural adaptation. Firestone and Scholl (2016) have recently addressed this issue in relation to top-down effects reported in perception research. Although the effect observed in the present study is not predicated on a “top-down effect” in the narrow sense, the same issue applies. The problem arises because although observers are told only to respond to the test stimuli, they are likely also considering the adapting stimuli in similar terms – i.e. the variability of the ensembles presented. They may become aware of the experimental manipulation and may modulate their responses in a way which then appears consistent with adaptation aftereffects, but in fact does not reflect a change in perception as a result of adaptation. It remains very difficult to truly separate decision bias from genuine perceptual bias in adaptation paradigms, however we think it is unlikely that our data are the product of decision bias and not a true adaptation aftereffect for two key reasons. Firstly, the observers were naïve to the purpose of the experiment and did not participate in more than one experiment. This is important because, to a naïve participant it would not be immediately obvious what the expected pattern of

responses would be – even if they identify that they are seeing high variability on one side during adaptation does that mean the experimenter is expecting they will be primed to see that side as more variable, or that they will adapt and require more variability on that side to equate the other? After the experiment each observer was asked what they thought the experiment was investigating and none gave a response which approximated the actual purpose of the study. The second reason that we are confident that the data do not reflect decision bias rather than perceptual bias is the nature of the task. Test slides were presented for a short period of time (500ms), during which the observer was required to compare the variability across two separate ensembles. Comparing two stimuli which are simultaneously visible is preferable to asking observers to make a judgment against an internal (e.g. rating scale or categorical judgment) or remembered standard (e.g. judgments across a delay interval). In our case the stimuli are both visible, facilitating a true perceptual judgment. If the results were to reflect decision bias rather than perceptual effects, the bias would need to be strong enough to result in responses which are contrary to the sensory perception of the observer (e.g. choosing left as more variable even though they can see the right is actually more variable), as would be necessary to produce a shift in the psychometric function of the magnitude we observed.

Conclusion

We found aftereffects of variance adaptation which generalise from colour to orientation. This suggests that visual variability is encoded by a domain-general mechanism which may serve to calibrate the visual system to the level of variability present in the observer's surroundings. Further research should seek to explore the limits of the generalisation from domain-to-domain, including whether transfer is equal in both directions (e.g. whether

variance adaptation aftereffects for colour can be observed from adaptation to orientation) and the possibility for adaptation to other summary statistical features (e.g. skew, kurtosis).

Context of the research

This study is part of an ongoing effort to understand how the brain encodes, and calibrates to, the complexity in the visual environment, particularly in the domain of colour. We have previously shown that observers are able to encode the mean of a colour array, for example (Maule & Franklin, 2016). This research was inspired by a study demonstrating adaptation to variance in orientation (Norman et al., 2015). This suggested that the visual system encodes variation “directly” as adaptation to one mean orientation (e.g. variation around vertical) would also affect judgments of the orthogonal orientation (e.g. variation around horizontal). Our study has shown that the same basic effect can be observed for colour. Many of the characteristics of visual summary statistical processing are common across visual domains (e.g. facial expression, size, brightness), which may be accounted for by a common neural substrate for encoding summary statistics, regardless of domain. Our final experiment supports this idea – showing that aftereffects of colour variance adaptation generalise to perception of orientation variance and is the first demonstration of such a cross-domain ensemble adaptation aftereffect.

Acknowledgments

This research was funded by the European Research Council (COLOURMIND: 772193).

References

- Allik, J., Toom, M., Raidvee, A., Averin, K., & Kreegipuu, K. (2014). Obligatory averaging in mean size perception. *Vision Research*, *101*, 34-40. doi:DOI 10.1016/j.visres.2014.05.003
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122-131. doi:DOI 10.1016/j.tics.2011.01.003
- Anobile, G., Burr, D. C., Iaia, M., Marinelli, C. V., Angelelli, P., & Turi, M. (2018). Independent adaptation mechanisms for numerosity and size perception provide evidence against a common sense of magnitude. *Sci Rep*, *8*(1), 13571. doi:10.1038/s41598-018-31893-6
- Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157-162. doi:10.1111/1467-9280.00327
- Ariely, D. (2008). Better than average? When can we say that subsampling of items is better than statistical summary representations? *Perception & Psychophysics*, *70*(7), 1325-1326. doi:10.3758/PP.70.7.1325
- 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), 1440-1449. doi:Doi 10.1037/A0036206
- Bauer, B. (2009). Does Stevens's Power Law for Brightness Extend to Perceptual Brightness Averaging? *Psychological Record*, *59*(2), 171-185.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433-436. doi:10.1163/156856897x00357
- Chetverikov, A., Campana, G., & Kristjansson, A. (2017). Representing Color Ensembles. *Psychological Science*, *28*(10), 1510-1517. doi:10.1177/0956797617713787
- Chong, S. C., Joo, S. J., Emmanouil, T. A., & Treisman, A. (2008). Statistical processing: not so implausible after all. *Perception & Psychophysics*, *70*(7), 1327-1334; discussion 1335-1326. doi:10.3758/PP.70.7.1327
- Corbett, J. E., & Melcher, D. (2014). Stable Statistical Representations Facilitate Visual Search. *Journal of Experimental Psychology-Human Perception and Performance*, *40*(5), 1915-1925. doi:Doi 10.1037/A0037375
- Corbett, J. E., & Oriet, C. (2011). The whole is indeed more than the sum of its parts: Perceptual averaging in the absence of individual item representation. *Acta Psychologica*, *138*(2), 289-301. doi:10.1016/j.actpsy.2011.08.002
- Corbett, J. E., Wurnitsch, N., Schwartz, A., & Whitney, D. (2012). An aftereffect of adaptation to mean size. *Visual Cognition*, *20*(2), 211-231. doi:Doi 10.1080/13506285.2012.657261
- Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America. A, Optics, image science, and vision*, *18*(5), 1016-1026.
- Dakin, S. C. (2012). Seeing statistical regularities. In J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organisation* (pp. 150-166). Oxford: Oxford University Press.
- Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, *37*(22), 3181-3192.
- de Fockert, J. W., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology*, *62*(9), 1716-1722. doi:10.1080/17470210902811249

- de Gardelle, V., & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(32), 13341-13346. doi:DOI 10.1073/pnas.1104517108
- Demeyere, N., Rzeskiewicz, A., Humphreys, K. A., & Humphreys, G. W. (2008). Automatic statistical processing of visual properties in simultanagnosia. *Neuropsychologia*, *46*(11), 2861-2864. doi:DOI 10.1016/j.neuropsychologia.2008.05.014
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781. doi:10.3389/fpsyg.2014.00781
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.
- Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*. doi:10.1017/S0140525X15000965
- 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), 432-446. doi:10.1037/xge0000053
- Haberman, J., Lee, P., & Whitney, D. (2015). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *J Vis*, *15*(4), 16. doi:10.1167/15.4.16
- Haberman, J., & Whitney, D. (2009). Seeing the Mean: Ensemble Coding for Sets of Faces. *Journal of Experimental Psychology-Human Perception and Performance*, *35*(3), 718-734. doi:Doi 10.1037/A0013899
- Haberman, J., & Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Attention Perception & Psychophysics*, *72*(7), 1825-1838. doi:Doi 10.3758/App.72.7.1825
- Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience*, *4*(2), 127-128. doi:10.1038/83945
- Im, H. Y., & Halberda, J. (2013). The effects of sampling and internal noise on the representation of ensemble average size. *Attention Perception & Psychophysics*, *75*(2), 278-286. doi:DOI 10.3758/s13414-012-0399-4
- Ishihara, S. (1973). *Ishihara's test chart for colour deficiency*. Tokyo: Kanehara Trading INC.
- JASP Team. (2018). JASP (Version 0.9).
- Kimura, E. (2018). Averaging colors of multicolor mosaics. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *35*(4), B43-B54. doi:10.1364/Josaa.35.000b43
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 14-14.
- Kuriki, I. (2004). Testing the possibility of average-color perception from multi-colored patterns. *Optical Review*, *11*(4), 249-257. doi:DOI 10.1007/s10043-004-0249-2
- Lanthon, P. (1998). *Album Tritan* (2nd ed.). Paris: Laboratoire de la Vision des Couleurs.
- Lanzoni, L., Melcher, D., Miceli, G., & Corbett, J. E. (2014). Global statistical regularities modulate the speed of visual search in patients with focal attentional deficits. *Frontiers in Psychology*, *5*(514), 1-12. doi:10.3389/fpsyg.2014.00514
- Lau, J. S., & Brady, T. F. (2018). Ensemble statistics accessed through proxies: Range heuristic and dependence on low-level properties in variability discrimination. *J Vis*, *18*(9), 3. doi:10.1167/18.9.3
- Leib, A. Y., Landau, A. N., Baek, Y., Chong, S. C., & Robertson, L. (2012). Extracting the mean size across the visual field in patients with mild, chronic unilateral neglect. *Frontiers in Human Neuroscience*, *6*. doi:10.3389/Fnhum.2012.00267

- Leib, A. Y., Puri, A. M., Fischer, J., Bentin, S., Whitney, D., & Robertson, L. (2012). Crowd perception in prosopagnosia. *Neuropsychologia*, *50*(7), 1698-1707. doi:10.1016/j.neuropsychologia.2012.03.026
- Li, A., & Lennie, P. (1997). Mechanisms underlying segmentation of colored textures. *Vision Research*, *37*(1), 83-97.
- Luo, A. X., & Zhou, G. (2018). Ensemble perception of facial attractiveness. *J Vis*, *18*(8), 7. doi:10.1167/18.8.7
- Marchant, A. P., & de Fockert, J. W. (2009). Priming by the mean representation of a set. *Quarterly Journal of Experimental Psychology*, *62*(10), 1889-1895. doi:10.1080/17470210902871045
- Marchant, A. P., Simons, D. J., & de Fockert, J. W. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta Psychologica*, *142*(2), 245-250. doi:DOI 10.1016/j.actpsy.2012.11.002
- Maule, J., & Franklin, A. (2015). Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of Vision*, *15*(4), 1-18. doi:10.1167/15.4.6
- Maule, J., & Franklin, A. (2016). Accurate rapid averaging of multihue ensembles is due to a limited capacity subsampling mechanism. *Journal of the Optical Society of America A*, *33*(3), A22-A29. doi:10.1364/JOSAA.33.000A22
- Maule, J., Stanworth, K., Pellicano, E., & Franklin, A. (2016). Ensemble perception of color in autistic adults. *Autism Research*, n/a-n/a. doi:10.1002/aur.1725
- 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. doi:Doi 10.1364/Josaa.31.000a93
- McCollough, C. (1965). Color Adaptation of Edge-Detectors in Human Visual System. *Science*, *149*(3688), 1115-1116. doi:10.1126/science.149.3688.1115
- Michael, E., de Gardelle, V., & Summerfield, C. (2014). Priming by the variability of visual information. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(21), 7873-7878. doi:DOI 10.1073/pnas.1308674111
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782-784.
- Morgan, M., Chubb, C., & Solomon, J. A. (2008). A 'dipper' function for texture discrimination based on orientation variance. *J Vis*, *8*(11), 9 1-8. doi:10.1167/8.11.9
- Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2015). Direct encoding of orientation variance in the visual system. *J Vis*, *15*(4), 3. doi:10.1167/15.4.3
- Oriet, C., & Brand, J. (2013). Size averaging of irrelevant stimuli cannot be prevented. *Vision Research*, *79*, 8-16. doi:DOI 10.1016/j.visres.2012.12.004
- 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. doi:Doi 10.1038/89532
- Pavlovskaya, M., Soroker, N., Bonneh, Y. S., & Hochstein, S. (2015). Computing an average when part of the population is not perceived. *J Cogn Neurosci*, *27*(7), 1397-1411. doi:10.1162/jocn_a_00791
- Payzan-LeNestour, E., Balleine, B. W., Berrada, T., & Pearson, J. (2016). Variance After-Effects Distort Risk Perception in Humans. *Current Biology*, *26*(11), 1500-1504. doi:10.1016/j.cub.2016.04.023
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437-442. doi:10.1163/156856897x00366
- Simons, D. J., & Myczek, K. (2008). Average size perception and the allure of a new mechanism. *Perception & Psychophysics*, *70*(7), 1335-1336. doi:Doi 10.3758/Pp.70.7.1335

- Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *J Vis*, *11*(12), 13. doi:10.1167/11.12.13
- Suarez-Pinilla, M., Seth, A. K., & Roseboom, W. (2018). Serial dependence in the perception of visual variance. *J Vis*, *18*(7), 4. doi:10.1167/18.7.4
- Sunaga, S., & Yamashita, Y. (2007). Global color impressions of multicolored textured patterns with equal unique hue elements. *Color Research and Application*, *32*(4), 267-277. doi:Doi 10.1002/Col.20330
- Sweeny, T. D., & Whitney, D. (2014). Perceiving Crowd Attention: Ensemble Perception of a Crowd's Gaze. *Psychological Science*, *25*(10), 1903-1913. doi:Doi 10.1177/0956797614544510
- Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2014). Ensemble perception of size in 4-5-year-old children. *Developmental Science*, *18*(4), 556-568. doi:10.1111/desc.12239
- Utochkin, I. S., & Tiurina, N. A. (2014). Parallel averaging of size is possible but range-limited: A reply to Marchant, Simons, and De Fockert. *Acta Psychologica*, *146*, 7-18. doi:DOI 10.1016/j.actpsy.2013.11.012
- Watamaniuk, S. N. J., & Duchon, A. (1992). The Human Visual-System Averages Speed Information. *Vision Research*, *32*(5), 931-941. doi:Doi 10.1016/0042-6989(92)90036-I
- Watamaniuk, S. N. J., & McKee, S. P. (1998). Simultaneous encoding of direction at a local and global scale. *Perception & Psychophysics*, *60*(2), 191-200.
- Webster, J., Kay, P., & Webster, M. A. (2014). Perceiving the average hue of color arrays. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *31*(4), A283-A292. doi:Doi 10.1364/Josaa.31.00a283
- Webster, M. A., De Valois, K. K., & Switkes, E. (1990). Orientation and spatial-frequency discrimination for luminance and chromatic gratings. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *7*(6), 1034-1049.
- Webster, M. A., & Mollon, J. D. (1994). The Influence of Contrast Adaptation on Color Appearance. *Vision Research*, *34*(15), 1993-2020. doi:Doi 10.1016/0042-6989(94)90028-0
- Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychon Bull Rev*, *19*(6), 1057-1064. doi:10.3758/s13423-012-0295-x
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology*, *69*, 105-129. doi:10.1146/annurev-psych-010416-044232
- Wolfe, B. A., Kosovicheva, A. A., Leib, A. Y., Wood, K., & Whitney, D. (2015). Foveal input is not required for perception of crowd facial expression. *J Vis*, *15*(4), 11. doi:10.1167/15.4.11
- Yamanashi Leib, A., Fischer, J., Liu, Y., Qiu, S., Robertson, L., & Whitney, D. (2014). Ensemble crowd perception: a viewpoint-invariant mechanism to represent average crowd identity. *J Vis*, *14*(8), 26. doi:10.1167/14.8.26
- Yang, Y., Tokita, M., & Ishiguchi, A. (2018). Is There a Common Summary Statistical Process for Representing the Mean and Variance? A Study Using Illustrations of Familiar Items. *Iperception*, *9*(1). doi:10.1177/2041669517747297
- Ying, H., Burns, E., Lin, X., & Xu, H. (2019). Ensemble statistics shape face adaptation and the cheerleader effect. *Journal of Experimental Psychology-General*, *148*(3), 421-436. doi:10.1037/xge0000564
- Ying, H., & Xu, H. (2017). Adaptation reveals that facial expression averaging occurs during rapid serial presentation. *J Vis*, *17*(1), 15. doi:10.1167/17.1.15

Zhaoping, L., & Zhe, L. (2012). Properties of V1 neurons tuned to conjunctions of visual features: application of the V1 saliency hypothesis to visual search behavior. *Plos One*, 7(6), e36223. doi:10.1371/journal.pone.0036223