The Quarterly Journal of Experimental Psychology

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

Visual cortex activation predicts visual preference: evidence from Britain and Egypt

Alexis D.J. Makin , Mai Helmy & Marco Bertamini

Publisher: Taylor & Francis & The Experimental Psychology Society
Journal: *The Quarterly Journal of Experimental Psychology*
DOI: 10.1080/17470218.2017.1350870

Visual cortex activation predicts visual preference: evidence from Britain and Egypt

Alexis D.J. Makin[1], Mai Helmy[2] and Marco Bertamini[1]

1: University of Liverpool
Department of Psychological Sciences
Eleanor Rathbone Building,
University of Liverpool,
Liverpool,
L69 7ZA
United Kingdom

2: Minoufiya University
Department of Psychology
Shibīn al Kawm,
Menoufia,
Egypt

Corresponding author
Alexis Makin – alexis.makin@liverpool.ac.uk
ORCID 0000-0002-4490-7400

Abstract

The term 'Perceptual goodness' refers to the strength, obviousness or salience of a visual configuration. Recent work has found strong agreement between theoretical, neural and behavioural measures of perceptual goodness across a wide range of different symmetrical visual patterns (Makin et al. 2016). We used these pattern types again to explore the relationship between perceptual goodness and aesthetic preference. A group of 50 UK participants rated the patterns on a 0-100 scale. Preference ratings positively correlated with four overlapping measures of perceptual goodness. We then replicated this finding in Egypt, suggesting that our results reflect universal aspects of human preference. The third experiment provided consistent results with a different stimulus set. We conclude that symmetry is an *aesthetic primitive* that is attractive because of the way it is processed by the visual system.

Keywords

Aesthetics; Holographic Model; Perceptual Goodness; Sustained Posterior Negativity; Symmetry

Introduction

Perceptual goodness

Perceptual goodness is a term from the early Gestalt school, referring to perceptual strength, obviousness or salience of a visual configuration (Koffka, 1935; Wertheimer, 1923, for a recent review see Wagemans, 2017). Reflectional symmetry is a *good gestalt*, where the structure is immediately apparent to human observers. Other symmetries, like repetition and rotation, are less salient (Mach, 1886). Different visual regularities are shown in Figure 1. Note how some of these are more obvious than others.

[Figure 1 about here]

There is no definitive list of rules governing perceptual goodness, although some have argued that it is linked to simplicity and redundancy (Attneave, 1954; Hochberg & McAlister, 1953; Pomerantz & Kubovy, 1986). Following this theme, Van der Helm and Leeuwenberg (1996) proposed their *holographic weight of evidence model*, which quantifies the perceptual goodness of different regular dot patterns. Their key formula states that $W = E/N$, where E is evidence for regularity and N is the total amount of information. For reflectional symmetry with a single fold, E is the number of mid-point collinear dot-pairs across the axis, and N is the total number of dots. Consequently, W is always 0.5, however many extra symmetrical pairs are added. For repetition, E is the number of repeated blocks minus 1, and N is again the number of dots. W goes down as we increase the number of dots, while W goes up if we increase the number of repeated blocks. For Glass patterns (Glass, 1969), E is the number of dipole dot pairs minus 1, and N is number of dots. W rapidly approaches 0.5 (the same as 1-Fold reflection) when the number of dipoles increases. The scope and assumptions of the holographic model have been debated (Olivers, Chater, & Watson, 2004; van der Helm & Leeuwenberg, 2004). However, the W goodness metric successfully predicts performance in regularity discrimination tasks (Nucci & Wagemans, 2007).

Makin et al. (2016) ran a series of 5 studies using regular patterns like those in Figure 1, where W ranged from 0.01 to 0.875 (and see Supplemental Material 1 Figure 1 for more examples, including the matched random patterns). W explained most variance in grand-average response time in their forced-choice regular/random discrimination tasks (r^2 = 0.88), while W also explained considerable variance in grand-average error rate (r^2 = 0.78).

Visual symmetry generates an Event Related Potential (ERP) component called the *Sustained Posterior Negativity* (SPN, Supplemental Material 1 Figure 2). Amplitude is more negative for symmetrical than random patterns, at posterior electrodes, from around 250 ms onwards (Bertamini & Makin, 2014; Makin, Wilton, Pecchinenda, & Bertamini, 2012; Norcia, Candy, Pettet, Vildavski, & Tyler, 2002). The SPN is generated by extrastriate visual regions including the Lateral Occipital Complex (LOC), where symmetry activations have also been detected with functional Magnetic Resonance Imaging (fMRI) (Kohler, Clarke, Yakovleva, Liu, & Norcia, 2016; Sasaki, Vanduffel, Knutsen, Tyler, & Tootell, 2005). Makin et al. (2016) found that the W strongly predicted grand-average SPN amplitude (r^2 = 0.86).

We thus have four separate estimates of perceptual goodness for the pattern in Figure 1. Two of these measures are behavioural (response time and error rate), one is neural (SPN amplitude) and one theoretical (W load). All the four measures are potentially limited (response time and error rate are noisy and distorted by ceiling effects, ERPs are noisy for different reasons, W is noiseless, but ignores some systematic influences on perceptual goodness). Nevertheless, these disparate measures of perceptual goodness are all strongly correlated (see scatterplots in Supplemental Material 1 Figure 3). This gives a solid foundation for our current work.

Scientific aesthetics

Early pioneers like Birkhoff (1932) and Eysenck (1941) proposed formula relating objective stimulus features to aesthetic preference. This research program continues (for review see Palmer, Schloss, & Sammartino, 2013) but faces substantial challenges. For one thing, it is impossible to evoke strong emotions like *intense fascination* (Kubovy, 2000) or *aesthetic rapture* (Markovic, 2012) with controlled stimuli on repeated trials. Furthermore, the human aesthetic faculty is tuned to gestalts, or wholes, and responses cannot be predicted by summing preference for parts (Holmes & Zanker, 2012). Some reliable effects have been discovered despite these limitations: Most people prefer symmetrical to random arrangements (Eisenman, 1967; Makin, Pecchinenda, & Bertamini, 2012) blue to brown (Palmer & Schloss, 2010) and smooth curvature to angularity (Bertamini, Palumbo, Gheorghes, & Galatsidas, 2016; Cotter et al., 2017).

The current project examined the relationship between perceptual goodness and visual preference. People may like the more obvious regularities, which have a high W load, and generate a large SPN. This prediction has a long history - many scholars have claimed that beauty arises from the balance or combination of two fundamental factors. The first factor can be described as '*order, unity or harmony*' and the second factor as '*complexity, multiplicity or diversity*' (Boselie & Leeuwenberg, 1985; Eisenman, 1967; Roberts, 2007). Classic work proposed that beauty = order X complexity (Eysenck, 1941) or that beauty = order / complexity (Birkhoff, 1932). In the holographic model, order is number of holographic identities (E) and complexity is the number of elements (N). Birkhoff's theory thus predicts that people should like high W patterns (because W = E/N).

More recent work also makes similar productions. The *fluency-attribution model* states that people are sensitive to the efficiency of their own perceptual and cognitive operations, and often like things that are processed fluently (Reber, 2012). The high W patterns are processed quickly and generate large neural response, and should theoretically be preferred.

Artists often exaggerate distinctive features to an unrealistic level (Ramachandran & Hirstein, 1999). It is essential for animals to look good to potential mates, and for flowers to look good to pollinating insects. This often means some form of phenotypic exaggeration, such as enlarged and brightly-coloured tail feathers, petals or other kinds of sexual dimorphism. Perhaps this applies more generally, and human aesthetic interest can be aroused by unusually high levels of visual excitation in the extrastriate symmetry network?

It is easy to find counter-claims. Resolution of perceptual ambiguity (Van de Cruys & Wagemans, 2011) or discovery of representational fit between visual and abstract layers (Palmer et al., 2013) can be more fundamental to artistic success than brute visual excitation. Second, some visual dimensions have an aesthetically optimal mid-point and unpleasant extremes (e.g. Berlyne, 1970; Redies, 2007; Spehar, Clifford, Newell, & Taylor, 2003). Finally, perfect symmetry may have a sterile rigidity about it, and many people may prefer imperfect symmetry disrupted by noise (Gartus & Leder, 2013; McManus, 2005).

Experiment 1 measured explicit preference for the pattern types in Figure 1. Whereas Makin et al. (2016) ran seperate studies on different groups of participants, here conditions were presented as five *sub-experiments, and all participants completed all five sub-experiments*. Will people prefer the high W patterns (that are discriminated efficiently and produce large brain response) or will they like less obvious regularities (that are discriminated less efficiently and produce an intermediate brain response)? Different accounts in scientific aesthetics make diverging predictions, so we had no strong a-priori hypothesis.

There were 100 participants in Experiment 1. The first group of 50 were English speakers recruited at the University of Liverpool, UK. The other 50 were Arab-speaking undergraduates at Minoufiya University in Egypt. Cross-cultural comparisons are important and useful, but the comparison between Western and Arab culture is particularly interesting. Classic Islamic art has a greater emphasis on abstract geometry (Gonzales, 2001), while classic Western art celebrates beauty in human faces and bodies. Soueif and Eysenck (1971) studied preferences for shape and found similar results in Britain and in Egypt; this has been confirmed recently (Bode, Helmy, & Bertamini, 2017). Although analysis of just two populations will never be sufficient to support claims of universality, it does at least go some way to addressing the generalizability of our results.

Experiment 1 Method

Fifty participants were from the University of Liverpool (mean age 28.2, min = 17, max = 58, 20 males, 4 left handed) and fifty were from Egypt (mean age 19.9, min = 19, max = 22, 3 males, 0 left handed). The Egyptian participants were younger that the UK sample on average, and there were a higher proportion of females. These confounds would be problematic if we were reporting large-cross cultural differences. However, results from UK and Egypt were similar.

Most participants were students or research staff. The study had local ethics committee approval and were conducted in accordance with the Declaration of Helsinki (revised 2008). In a previous unpublished experiment on the same topic (but without neat matching to the 5 studies of Makin et al. (2016), we found strong effects with 28 participants (minimum $r^2 = 0.69$). We were confident that 100 was an adequate sample size.

Each participant was involved in 5 short *sub-experiments*, matched to the 5 studies reported in Makin et al. (2016). Patterns types were identical to Makin et al. (2016) and like those shown in Figure 1. Individual patterns were generated afresh on each trial according to an algorithm programmed in Python using PsychoPy (Peirce, 2007) (see open science framework for source code and raw data https://osf.io/3amna/). No participant ever saw the same exact pattern twice, and patterns were never repeated across participants. On each trial, a pattern was shown for 3 seconds then participants entered a preference judgment on a 0-100 Likert scale. The text above the scale read 'How much did you like that pattern?' and the extremes of the scale were labelled '0= not at all' and 'extremely = 100' (Supplemental Figure 4).

In each sub-experiment, different examples of the regular patterns were presented 6 times, and the number of random patterns was balanced to match the total number of regular ones. In other words, half the trials were regular and half were random. This was consistent with the design of the ERP and behavioural experiments reported in Makin et al. (2016). Participants were not restrained in a chin rest, but pattern size was approximately 5 X 5° visual angle (as in Makin et al. 2016). The order of the 5 experiments varied between participants, and the experiment lasted around 25 minutes in total.

Relative preference was computed in each sub-experiment and for each participant. This involved three steps: 1) mean preference for the six regular patterns in each condition was computed, 2) mean rating for all random patterns was computed, 3) relative preference score was computed as the regular mean – random mean. Positive values indicate relative preference for regularity. Relative preference is a superior metric to absolute rating because it isolates aspects of preference exclusively related to W.

The correlation between grand-average relative preferences and W was assessed with Pearson's r. However, the holographic model does not stipulate how to calculate W for anti-symmetry like that used in sub-experiments 4 and 5 (see van der Helm & Treder 2009). We excluded these data points, so the W vs. relative preference correlations were based on the other 16 conditions. The correlation between relative preference and 3 measures of perceptual goodness from Makin et al. (2016) was similar. Here we simply use the grand-average response times, error rate and SPN amplitudes (although these were obtained from different groups of participants and conditions) as our predictor variables. These correlations were based on 18 data points, because we have these measures for anti-symmetry. SPN amplitude was defined as regular-random at PO7/8 electrodes, from 300-400 ms post stimuli. This was chosen because the correlation between SPN amplitude and W peaked in this early window, although it was very strong throughout the traditional 300-1000 ms SPN interval.

Relative preference in the sub-experiments was analysed with repeated measures ANOVA. The Greenhouse-Geisser correction factor was applied whenever the assumption of sphericity was violated (Mauchly's W, $p < 0.05$).

Experiment 1 Results

We first examined *relative preference* scores across all 5 sub-experiments. Note that relative preference is calculated for each participant and condition. It is the difference between the mean rating for a regular pattern and the mean rating from the matched random pattern. Positive values indicate preference for regularity. Figure 2 shows that there was a strong linear relationship between W and grand-average relative preferences ($r = 0.898$, $p < 0.001$). Relative preference was also correlated with SPN (r -0.820, $p < 0.001$), response time (r = -0.758, $p < 0.001$) and error rate ($r = 0.739$, $p < 0.001$).

[Figure 2 about here]

Figure 5 shows r^2 values for UK and Egyptian samples alongside each scatter-plot. This gives the proportion of variance in *grand-average* relative preferences explained by different measures of perceptual goodness. Our goodness metrics explained more variance in UK than Egypt. However, we were mindful that analysis based on aggregated data leads to an overestimation of effect size (Brand & Bradley, 2012). We thus ran three additional analyses that take participant and trial level into account.

First, we ran an equivalent analysis on all individual participants. Each participant provides 18 relative preference scores. We then analysed this in the same way as the grand-averages. Figure 2A shows r^2 values from each participant organized cumulatively (with rare correlations with atypical sign coded as $r^2 = 0$). This graphical convention means that area below the line gives the proportion of variance explained by the predictor (and therefore the areas above the line is unexplained variance). Figure 2A shows this for the UK (blue) and Egyptian samples (yellow). Area under the line (variance explained) ranged between 0.33 and 0.54.

Second, we computed regression coefficients and p-values with Linear Mixed Effects Analysis, which uses data from all trials and participants (LME4 library in RLME4 library in R, Bates, Mächler, Bolker, & Walker, 2015). Relative preference was the dependent variable. Country, Sub-Experiment, Participant and Trial were included as random factors. This showed than an increase in W from 0 to 1 predicts a 44.45 point increase in relative preference (Relative preference = 44.45W + 2.97; $\chi 2$ (1) = 1479.4, $p < 0.001$). Strong effects were also found for SPN (Relative preference = -10.01 SPN amplitude(μV) + 7.18; $\chi 2$ (1) = 1226.1, $p < 0.001$), Response time (Relative preference = -33.99 RT(s) + 53.60, $\chi 2$ (1) = 1098.8, $p < 0.001$) and Error rate (Relative preference = -78.44 Error rate + 31.678; $\chi 2$ (1) = 1373.5, $p < 0.001$).

[Figure 3 about here]

So far, the results fit a simple story of *Relative preference = Perceptual Goodness* and UK = Egypt. This was generally true when each of the 5 sub-experiments were analysed independently, however, there were a few systematic exceptions. Relative preference scores are shown in Figure 3B – F.

In the sub-experiment 1, there were main effects of and Regularity (F (1,98) = 167.690, $p < 0.001$, partial η^2 = 0.631) and N-Dots (F (1.383,135.522) = 47.077, $p < 0.001$, partial η^2 = 0.324), but no Regularity X N dots interaction (F (1.827,75.179.063) = 1.596, p = 0.207, Figure 3B). There would be an interaction here if relative preferences were perfectly predicted by W, because N-Dots has a unique effect on repetition according to the holographic model (van der Helm & Leeuwenberg, 1996). There was also a weak N-Dots X Country interaction (1.383,135.522) = 3.543, p = 0.048, partial η^2 = 0.035), because the effect of N-dots was slightly weaker in the UK (F (1.214, 59.501) = 19.376, $p < 0.001$ partial η^2 = 0.283) than Egypt (F (1.432, 70.178) = 28.235, $p < 0.001$, partial η^2 = 0.366). There was no main effect of Country (F (1,98) < 1)

The preferences in sub-experiment 2 were closely related to W (Figure 3C). Participants liked reflection and Glass patterns almost equally, and liked both far more than repetition (F (2,196) = 55.102, $p < 0.001$, η^2 = 0.360). There was no main effect of Country (F (1,98) = 2.632, p = 0.108) and no Regularity X Country interaction (F (2,196) = 2.395, p = 0.094).

According to the holographic model, there is a general increase in W with the number of axes of reflection, but 3 and 5-Fold reflections have slightly lower W-loads than 2 and 4-Fold reflections (van der Helm, 2011). The dip at 3 and 5-Folds was not apparent in sub-experiment 3, although there was not increase either (Figure 2D). Relative preference generally increased with Folds (F (2.934, 287.526) = 110.582, $p < 0.001$, $\eta^2 = 0.530$), and preferences were higher in Egypt than the UK (F (1,98) = 14.508, $p < 0.001$ partial $\eta^2 =$ 0.125). There was a Folds X Country interaction (F (2.934, 287.526) = 11.486, $p < 0.001$, partial $\eta^2 = 0.105$) because the effect of Folds was weaker in the UK (F (2.709, 132.721) = 36.483, $p < 0.001$, partial $\eta^2 = 0.427$) than Egypt (F (2.861, 140.208) = 75.134, $p < 0.001$, partial $\eta^2 = 0.605$).

In anti-symmetry patterns, black elements are paired with white, and white elements are paired with black. Our sub-experiments 4 and 5 measured preference for symmetry and anti-symmetry with 1 or 4-Folds (Figure 3E and F). The holographic model is silent about the goodness of anti-symmetry (because it is not considered to be a basic visual regularity, see van der Helm & Treder 2009). However, we know that anti-symmetry is not discriminated as efficiently as symmetry under many conditions (e.g. Mancini, Sally, & Gurnsey, 2005) and sometimes produces a slightly smaller SPN (Makin et al., 2016). Results for these sub-experiments are shown in Figure 2E and F. It is instructive to analyse these together. The participants liked 4-Fold patterns more than 1-Fold patterns (F (1,98) = 167.783, $p < 0.001$, partial $\eta^2 = 0.631$) and symmetry more than anti-symmetry (F (1,98) = 102.316, $p < 0.001$, partial $\eta^2 = 0.511$). There was no main effect of Country (F (1,98) = 2.681, $p = 0.105$). However, there was a Folds X Country interaction (F (1,98) = 16.748, $p <$ 0.001, partial $\eta^2 = 0.146$) and a three-way interaction between Folds, Regularity and Country (F (1,98) = 8.891, $p = 0.004$, partial $\eta^2 = 0.083$). For 1-Fold patterns, there was no

Regularity X Country interaction ($F(1,98) < 1$), and no difference between UK and Egypt ($F(1, 98) < 1$). Conversely, for 4-fold patterns, there was both a Regularity X Country interaction ($F(1,98) = 7.564$, $p = 0.007$, partial $\eta^2 = 0.072$) and higher preferences overall in Egypt ($F(1,98) = 9.722$, $p = 0.002$, partial $\eta^2 = 0.090$).

Finally, we note that nearly every regularity was preferred to the matched random patterns presented in the sub-experiment. In other words, relative preference was always > 0 (one sample t tests, $p < 0.003$), with the sole exception of 20-dot repetition in sub-Experiment 1 ($t(99) = -1.196$, $p = 0.235$).

Experiment 1 Discussion

Preference for different types of regularity increased with perceptual goodness (whether goodness is operationalized as W, SPN amplitude, response time or error rate). Participants thus liked the most obvious regularities, which produced the largest SPN. This supports the idea that Beauty = Order/Complexity (Birkhoff, 1932). Results are also consistent with the fluency account of aesthetics (Reber, 2012) and with the idea that people are attracted to visual exaggerations (Ramachandran & Hirstein, 1999). Our results suggest that symmetry is an *aesthetic primitive*: A visual property that is inherently interesting because of the way it is processed by the visual system (Latto, Brain, & Kelly, 2000). Although there were a few minor differences, results from the UK sample were closely replicated in Egypt.

An excessively bold conclusion here would be *'Attraction to high W symmetry is a universal law of aesthetics, innately hardwired into visual and emotional brain areas'*. Of course, it would be naïve to make such a strong claim based on the available evidence. We therefore propose a nuanced version. Humans are hardwired so that some perceptual abilities always emerge whenever infants grow up in a typical visual environment. Through some combination of innate preparedness and exposure, the adult extrastriate cortex becomes better at discriminating some visual regularities than others (the holographic model provides a good estimate of adult sensitivity to different regularities). Preference ratings are linked to this regularity-sensitivity when regularity is the most prominent dimension in the data set. In the UK and Egypt, most students prefer the more salient regularities. This may generalise to most humans. More detailed proposals about evolution and development of symmetry perception and preference are included in van der Helm (2011).

Although the UK and Egyptian datasets were mostly similar, Egyptian participants tended to give higher ratings to the multiple symmetries. This might be because abstract geometry and multiple symmetries are celebrated in Islamic art (Gonzales, 2001). Such effects could be superimposed on other universal laws of aesthetics (see Carbon et al. 2010 for analysis of the dynamics of fashion). However, there are other possibilities here as well. Perhaps everyone liked the multiple symmetries, but the UK sample were less inclined to use the high end of the response scale? Such limitations of the Likert scale procedure make us cautious about over-interpreting between-subject effects.

Precise predictions of the holographic model could be tested in future work. For instance, preference should increase with the number of repeated blocks in a repetition pattern, and when the elements in a repetition can be grouped into salient blobs (Csathó, van der Vloed & van der Helm, 2003).

Furthermore, the holographic model (van der Helm & Leeuwenberg, 1996) can be contrasted with the alternative *transformational model* (Gardner 1974). Amongst other things, these two models make different claims about the goodness of the N-Fold reflections used in sub-Experiment 3. Van der Helm (2011) quantified the predictions of the transformational model (which we denote T) as 1/(1-2F); where F is the number of folds. This means that T increases monotonically with F, and asymptotes at high F. In contrast, the relationship between W and F is similar but *non-monotonic*, with a 'goodness dip' at 3 and 5-Fold (Figure 1). Makin et al. (2016) tested whether W (holographic model) or T (transformational model) explained more variance in their DVs. T explained more variance in RT and error rate (although this can be explained by ceiling effects when F => 2). Conversely, W explained more variance in early SPN amplitude, and the predicted dip at 3 and 5-Fold was apparent. Previously, Wenderoth and Welsh (1998) found that discrimination of 3-Fold symmetry was often very like 2-Fold symmetry discrimination (and sometimes worse). Our sub-experiment 3 found that relative preference for 2-Fold was approximately equal to 3-Fold, and relative preference for 4-Fold was approximately equal to 5-Fold. These results lie roughly between the predictions W and T (rather like the results of Wenderoth & Welsh, 1998).

We should also consider the generality of the sub-experiment 3 results. All patterns had 1 vertical axis, 2-Fold and 4-Fold reflections had additional horizontal axes, while 3, 4 and 5-Fold reflections had additional diagonal axes. The W scores for each of the n-Fold reflections would remain the same if they were globally rotated and all vertical and horizontal axes were eliminated. However, vertical and horizontal orientations might have a special status for the visual system (e.g. Wenderoth, 1994) so global rotation could potentially modulate preference independently of W. Indeed, it is certainly true that not all reliable goodness differences are captured by W differences. These anomalies can sometimes be plausibly explained secondary influences on early visual processing (e.g. Csathó et al. 2003). This is also a topic for future work.

Experiment 2

So far it seems that preferences for abstract symmetrical patterns is linearly related to perceptual goodness, as defined by the holographic model. However, our first experiment did not test this directly. Our low and high W patterns were built from different elements, and tested in different sub-experiments (Figure 1). In Experiment 2, we measured preference for symmetry + noise patterns which spanned the full W range. Example patterns are shown in Figure 4A. The proportion of symmetry varied between 0 and 100% in 20% increments (We called this proportion 'Psymm'). There is a linear relationship between W and Psymm. These kinds of patterns were used in an ERP study by Palumbo, Bertamini and Makin (2015), who found that SPN amplitude also scaled with Psymm (Supplemental Material 1 Figure 5), while previous work has shown similar scaling of the extrastriate BOLD response and behavioural performance (Sasaki et al., 2005, Barlow & Reeves, 1979, and see van der Helm 2010 for theoretical analysis). We thus expected that relative preference would increase with Psymm in Experiment 2.

Experiment 2 Method

Twenty-four participants from the University of Liverpool were involved in Experiment 2 (aged 18 to 35, mean age 21.6, 9 Male, 2 Left-Handed). The preference rating protocol was the same as that used in Experiment 1. There were 10 repeats of each Psymm level, and 50 random exemplars, giving 100 trials in total. The stimulus generation algorithm for these patterns is described by Palumbo et al. (2015) and considerations about W calculation are in Supplemental Material 1 of Makin et al. (2016). To illustrate, consider that an 80% symmetry is one where 80% of the dots are positioned symmetrically on a grid, and the other 20% are positioned randomly (but might form accidental pairs across 1 or more of the axis). W was estimated by multiplying P symm by W for a 4-Fold reflection (0.875) by P Symm (e.g. 0.875 X 0.8 = 0.7) and ignoring accidental pairing.

Experiment 2 Results and Discussion

As expected, relative preference increased with Psymm ($F (2.039, 46.893) = 142.994$, $p < 0.001$, $\eta^2 = 0.861$), and the difference between each increment was significant ($p < 0.002$, Figure 4B). The 5 grand-average relative preference scores were almost perfectly correlated with W ($r = 0.98$, $p = 0.004$) and grand-average SPN amplitude ($r = 0.99$, $p = 0.001$). At an individual participant level, W explained 84% of variance in relative preference (Figure 4C) and SPN amplitude explained 88% of variance in relative preference (Figure 4D).

The results of Experiment 2 were thus very simple. Participants liked perfect symmetry, and preference increased linearly with Psymm. These results mirror other recent work by Gartus and Leder (2014), who also found that most participants preferred perfect symmetries to broken symmetries with slight imperfections.

[Figure 4 about here]

General Discussion

Makin et al. (2016) found strong agreement between four different measures of perceptual goodness (W, SPN amplitude, response time and error rate). In Experiment 1, we confirmed that participants gave higher preference ratings to the higher W patterns which produce larger SPNs. These effects were replicated in UK and in Egypt. Experiment 2 found that participants liked perfect regularity, which produces the largest SPN (and not regularity with a degree of noise).

We found that the relationship between goodness and preference is linear. There was no evidence that people liked mid-level goodness, where the structure was not too obvious. At least for the patterns used here, we can confidently assert that preference is a straight function of perceptual goodness. Preference for high W patterns is consistent an aesthetic formula which states that beauty = order / complexity (Birkhoff, 1932). Given these results, we suggest that symmetry is an aesthetic primitive, that is, a feature whose appeal derives directly from the way it is processed by the visual system (Latto et al., 2000).

Although we aimed to examine divergent accounts, we are mindful that our experiments cannot definitively falsify them. In scientific aesthetics, we must always be cautious about generalizing claims far beyond the tested stimuli. Future experiments using other sets (e.g. real objects, faces or art) might discover a different relationship between symmetry and preference. There may indeed be cases where perfect symmetry looks sterile and rigid (McManus, 2005) and other cases where it is preferred (Gartus and Leder, 2013). Indeed, there is some evidence that symmetry preferences are not uniform across categories (Little, 2014). For instance, painting or film where every visual detail was arranged symmetrically would look obviously wrong, even though compositional balance is often desirable (Arnheim, 1974). This kind of 'gestalt nightmare' was analysed at length in Makin (2017).

Before moving on, we must consider one alternative explanation for our results. We assume that preference reports were strongly influenced by the perceptual goodness (perceptual goodness > preference > report). However, participants probably recognized that their own preferences were determined by perceptual goodness. They might have then taken a cognitive short-cut, and reported perceptual goodness directly, rather than dwelling on how much they liked the patterns (perceptual goodness > report). Did our participants bypass their aesthetic faculties altogether, and report perceptual goodness directly (perhaps mentally relabelling the response scale)? We cannot completely rule this out. However, previous studies have found a strong relationship between the salience of symmetry and implicit preferences (Makin et al. 2012, Bertamini et al. 2013), so it is unlikely that pure preference ratings (without the cognitive short-cut) would be completely different.

Many animals, including insects, fish and birds, use phenotypic symmetry in mate selection and food choice, and humans often use symmetry to judge sexual attractiveness (Grammer, Fink, Møller, & Thornhill, 2003). However, we suggest that preference for abstract symmetry is NOT merely an overgeneralization of innate mate selection strategies. Consider that multiple-axes symmetry is not face-like or body-like, but our participants liked it more than single-axis symmetry. Furthermore, Glass patterns are not at all face-like, but these were liked nearly as much as reflection. In summary, W is a far better predictor of preference than biological relevance. We propose that symmetry detection has a broad perceptual utility, and the appeal of symmetry is directly related to the strength of the symmetry signal in the extrastriate symmetry network.

Does this mean that other aesthetic accounts that emphasise ambiguity and resolution (Van de Cruys & Wagemans, 2011), representational fit (Palmer et al., 2013) and imperfection (McManus, 2005) are wrong? We do not go that far: Instead, the results force us to think about the scope of different ideas in scientific aesthetics. Perhaps preference is directly linked to perceptual goodness when the stimuli are tightly controlled, and when patterns are presented in a quasi-psychophysical lab experiment. Other accounts may describe the psychodynamics of aesthetic experience in the real world, where stimuli are multi-dimensional, and aesthetic evaluation is optional and unconstrained. Even though scientific aesthetics is an old enterprise, it remains at an early stage of development. This kind of distinction is vital if we are to apply our theoretical insights correctly.

Acknowledgements

This work was partly sponsored by the ESRC (ES/K000187/1) and Leverhulme Trust (ECF–721-2012).

References

Arnheim, R. (1974). *Art and visual perception: A psychology of the creative eye*. Berkley: University of California Press.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*(3), 183–193. DOI: 10.1037/h0054663

Barlow, H. B., & Reeves, B. C. (1979). Versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Research*, *19*(7), 783–793. DOI: 10.1016/0042-6989(79)90154-8

Berlyne, D. E. (1970). Novelty, complexity and hedonic value. *Perception and Psychophysics*, *8*, 279–286. DOI: 10.3758/BF03212593

Bertamini, M., Makin, A.D.J. & Rampone, G. (2013). Implicit association of symmetry with positive valence, high arousal and simplicity. *I perception*, 4, 317–327.

Bertamini, M., & Makin, A. D. J. (2014). Brain Activity in Response to Visual Symmetry. *Symmetry*, *6*, 975–996. DOI: 10.3390/sym6040975

Bertamini, M., Palumbo, L., Gheorghes, T. N., & Galatsidas, M. (2016). Do observers like curvature or do they dislike angularity? *British Journal of Psychology*, *107*(1), 154–178. DOI: 10.1111/bjop.12132

Birkhoff, G. D. (1932). *Aesthetic Measure*. Cambridge, Mass: Harvard Press.

Bode, C., Helmy, M., & Bertamini, M. (2017). A cross-cultural comparison for preference for symmetry: comparing British and Egyptians non-experts.

Boselie, F., & Leeuwenberg, E. (1985). Birkhoff Revisited: Beauty as a Function of Effect and Means. The American Journal of Psychology, 98(1), 1–39. DOI: 10.2307/1422765

Brand, A., & Bradley, M. T. (2012). More Voodoo Correlations: When Average-Based Measures Inflate Correlations. *The Journal of General Psychology*, *139*(4), 260–272. DOI: 10.1080/00221309.2012.703711

Carbon, C.C. (2010)The cycle of preference: Long-term dynamics of aesthetic appreciation. Acta Psychologia. 134, 233–244. DOI: 10.1016/j.actpsy.2010.02.004.

Cotter, K.N. Silvia, P.J., Bertamini, M., Palumbo, L., Vartanian, O. (2017). Curve Appeal: Exploring Individual Differences in Preference for Curved Versus Angular Objects. *i-Perception, 8*, 2.

Csathó, Á., van der Vloed, G., & van der Helm, P. A. (2003). Blobs strengthen repetition but weaken symmetry. *Vision Research*, *43*(9), 993–1007. DOI: 10.1016/S0042-6989(03)00073-7

Eisenman, R. (1967). Complexity-Simplicity: I. Preference For Symmetry And Rejection Of Complexity. *Psychonomic science, 8(4)*, 169-170.

Eysenck, H. (1941). The empirical determination of an aesthetic formula. *Psychological Review*, *48*, 83–92. DOI: 10.1037/h0062483

Garner, W. R. (1974). *The processing of information and structure*. Pontomac, MD: Earlbaum.

Gartus, A. & Leder, H. (2013). The small step toward asymmetry: aesthetic judgment of broken symmetries. i-Perception, 4(5), 361-364.

Glass, L. (1969). Moire effect from random dots. *Nature*, *223*, 578–580.

Gonzalez, V. (2001). *Beauty and Islam: Aesthetics in Islamic art and architecture*. London: IB Tauris.

Grammer, K., Fink, B., Møller, A. P., & Thornhill, R. (2003). Darwinian aesthetics: sexual selection and the biology of beauty. *Biological Reviews*, *78*(3), 385–407. DOI:

10.1017/s1464793102006085

Hochberg, J., & McAlister, E. (1953). A Quatitative approach to figural goodness. *Journal of Experimental Psychology*, *46*(5), 361–364. DOI: 10.1037/h0055809

Holmes, T., & Zanker, J. M. (2012). Using an oculomotor signature as an indicator of aesthetic preference. *I-Perception*, *3*(7), 426–439. DOI: 10.1068/i0448aap

Koffka. (1935). *Priciples of Gestalt Psychology*. New York: Harcourt, Brace & World.

Kohler, P. J., Clarke, A., Yakovleva, A., Liu, Y., & Norcia, A. M. (2016). Representation of Maximally Regular Textures in Human Visual Cortex. *The Journal of Neuroscience*, *36*(3), 714–29. DOI: 10.1523/JNEUROSCI.2962-15.2016

Kubovy, M. (2000). Visual Aesthetics. In K. A.E (Ed.), *Encyclopedia of Psychology*. New York: Oxford University Press.

Latto, R., Brain, D., & Kelly, B. (2000). An oblique effect in aesthetics: Homage to Mondrian (1872 - 1944). *Perception*, *29*(8), 981–987. DOI: 10.1068/p2352

Little, A. (2014). Domain Specificity in Human Symmetry Preferences: Symmetry is Most Pleasant When Looking at Human Faces. *Symmetry*, 6(2), 222–233. DOI: 10.3390/sym6020222

Mach, E. (1886). *The analysis of sensations and the relation of the psysical to the psychical*. New York: Dover.

Makin, A.D.J. (2017). The Gap Between Aesthetic Science and Aesthetic Experience. *Journal of conciousness studies* (1-2) 184–213.

Makin, A. D. J., Pecchinenda, A., & Bertamini, M. (2012). Implicit affective evaluation of visual symmetry. *Emotion*, *12*(5), 1021–1030. DOI: 10.1037/a00269241

Makin, A. D. J., Wilton, M. M., Pecchinenda, A., & Bertamini, M. (2012). Symmetry perception and affective responses: A combined EEG/EMG study. *Neuropsychologia*, *50*(14), 3250–3261. DOI: 10.1016/j.neuropsychologia.2012.10.003

Makin, A. D. J., Wright, D., Rampone, G., Palumbo, L., Guest, M., Sheehan, R., … Bertamini, M. (2016). An electrophysiological index of perceptual goodness. *Cerebral Cortex*, *26, 4416-4434.* DOI: 10.1093/cercor/bhw255

Mancini, S., Sally, S. L., & Gurnsey, R. (2005). Detection of symmetry and anti-symmetry. *Vision Research*, *45*(16), 2145–2160. DOI: 10.1016/j.visres.2005.02.004

Markovic, S. (2012). Components of aesthetic experience: aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *I Perception*, *3*(1), 1–17. DOI: 10.1068/i0450aap

McManus, C. (2005). Symmetry and asymmetry in art and aesthetics. *European Review*, *13*(157–180).

Norcia, A. M., Candy, T. R., Pettet, M. W., Vildavski, V. Y., & Tyler, C. W. (2002). Temporal dynamics of the human response to symmetry. *Journal of Vision*, *2*(2), 132–139. DOI: 10.1167/2.2.1

Nucci, M., & Wagemans, J. (2007). Goodness of regularity in dot patterns: global symmetry, local symmetry, and their interactions. *Perception*, *36*(9), 1305–1319. DOI: doi.org/10.1068/p5794

Olivers, C. N. L., Chater, N., & Watson, D. G. (2004). Holography does not account for goodness: A critique of van der Helm and Leeuwenberg (1996). *Psychological Review*, *111*(1), 242–260. DOI: 10.1037/0033-295x.111.1.242

Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(19), 8877–8882. DOI: 10.1073/pnas.0906172107

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, *64*, 77–107. DOI: 10.1146/annurev-psych-120710-100504

Palumbo, L., Bertamini, M., & Makin, A. (2015). Scaling of the extrastriate neural response to symmetry. *Vision Research*, *117*, 1–8. DOI: 10.1016/j.visres.2015.10.002

Pomerantz, J.R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization: Simplicity and likelihood principles. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), Handbook of perception and human performance (pp. 36: 1-45). New York: Wiley.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. DOI: 10.1016/j.jneumeth.2006.11.017

Ramachandran, V. S., & Hirstein, W. (1999). The science of art: a neurological theory of aesthetic experience. *Journal of Conciousness Studies*, *6*(6–7), 15–31.

Reber, R. (2012). Processing fluency, aesthetic pleasure, and culturally shared taste. In A. P. Shimamura & S. E. Palmer (Eds.), *Aesthetic science: Connecting mind, brain, and experience* (pp. 223–249). New York: Oxford University Press.

Redies, C. (2007). A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision*, *21*(1), 97–117. DOI: 10.1163/156856807782753886

Roberts, M. N. (2007). Complexity and aesthetic preference for diverse visual stimuli (PhD Thesis) Universitat de les Illes Balears, Spain.

Sasaki, Y., Vanduffel, W., Knutsen, T., Tyler, C. W., & Tootell, R. (2005). Symmetry activates extrastriate visual cortex in human and nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 3159–3163. DOI:

10.1073/pnas.0500319102

Soueif M.I., & Eysenck, H.J. (1971). Cultural differences in aesthetic judgment. *International Journal of Psychology*. 6*(4)*, 293-198.

Spehar, B., Clifford, C. W. G., Newell, B. R., & Taylor, R. P. (2003). Universal aesthetic of fractals. *Computers & Graphics-Uk*, *27*(5), 813–820. DOI: 10.1016/s0097-8493(03)00154-7

Van de Cruys, S., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *I Perception*, *2*(9), 1035–1062. DOI: 10.1068/i0466aap

van der Helm, P. A. (2010). Weber-Fechner behavior in symmetry perception? *Attention Perception & Psychophysics*, *72*(7), 1854–1864. DOI: 10.3758/app.72.7.1854

van der Helm, P. A. (2011). The influence of perception on the distribution of multiple symmetries in nature and art. *Symmetry*, *3*(1), 54–71. DOI: 10.3390/sym3010054

van der Helm, P. A., & Leeuwenberg, E. L. J. (1996). Goodness of visual regularities: A nontransformational approach. *Psychological Review*, *103*(3), 429–456. DOI: 10.1037/0033-295x.103.3.429

van der Helm, P. A., & Leeuwenberg, E. L. J. (2004). Holographic goodness is not that bad: Reply to Olivers, Chater, and Watson (2004). *Psychological Review*, *111*(1), 261–273.

DOI: 10.1037/0033-295x.111.1.261

van der Helm, P. A., & Treder, M. S. (2009). Detection of (anti)symmetry and (anti)repetition: Perceptual mechanisms versus cognitive strategies. Vision Research, 49(23), 2754–2763. DO:j.visres.2009.08.015

Wagemans, J. (2017). Perceptual Organization. In J.T. Wixted (Series Ed.) & J. Serences (Vol.Ed.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience: Vol. 2. Sensation, Perception & Attention*. Hoboken, NJ: John Wiley.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II, in Psycologische Forschung, 4:301-350. Translation published in Ellis W. 1938. A source book of Gestalt psychology. (pp. 71–88). London: Routledge & Kegan Paul.

Wenderoth, P. (1994). The salience of vertical symmetry. *Perception*, *23*(2), 221–236. DOI: 10.1068/p230221

Wenderoth, P., & Welsh, S. (1998). Effects of pattern orientation and number of symmetry axes on the detection of mirror symmetry in dot and solid patterns. *Perception*, *27*(8), 965–976. http://doi.org/10.1068/p270965

Figure Legends

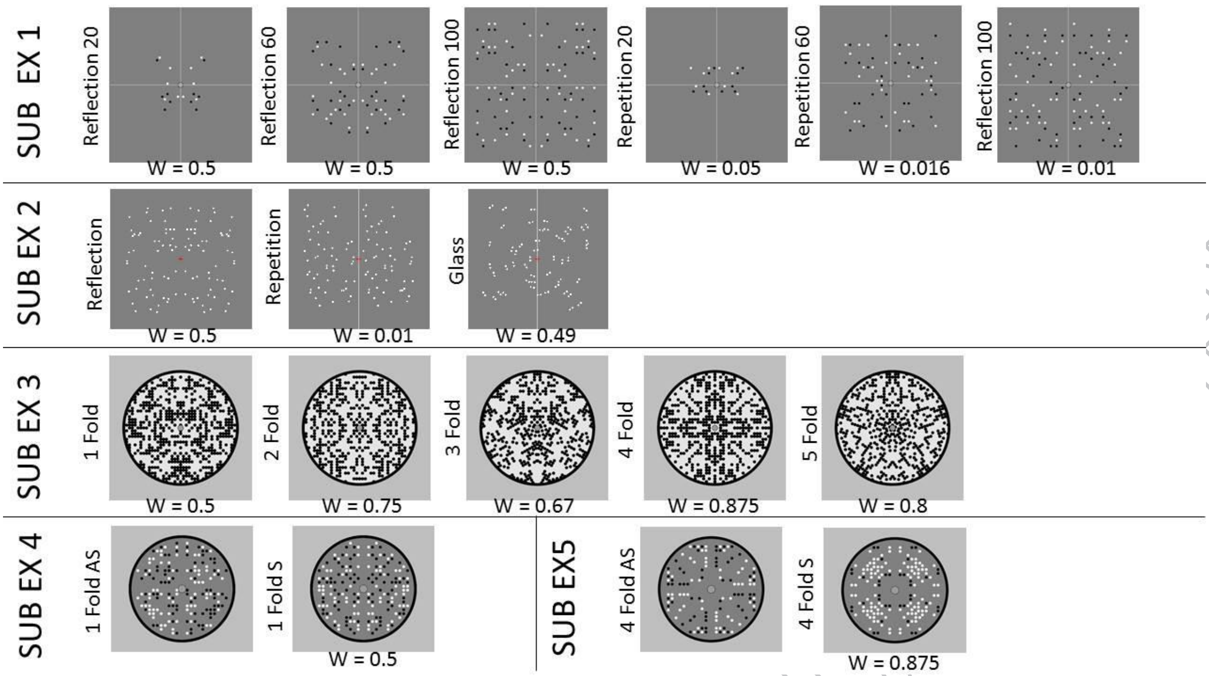
Figure 1. Examples of the regular patterns from our five sub-experiments. Participants completed all five experiments. Random patterns are not shown here, but available in Supplemental Material 1 Figure 1. These are just examples. Different patterns were generated afresh on every trial, so no participant ever saw the exact same stimulus set. The same pattern types were used by Makin et al. (2016).

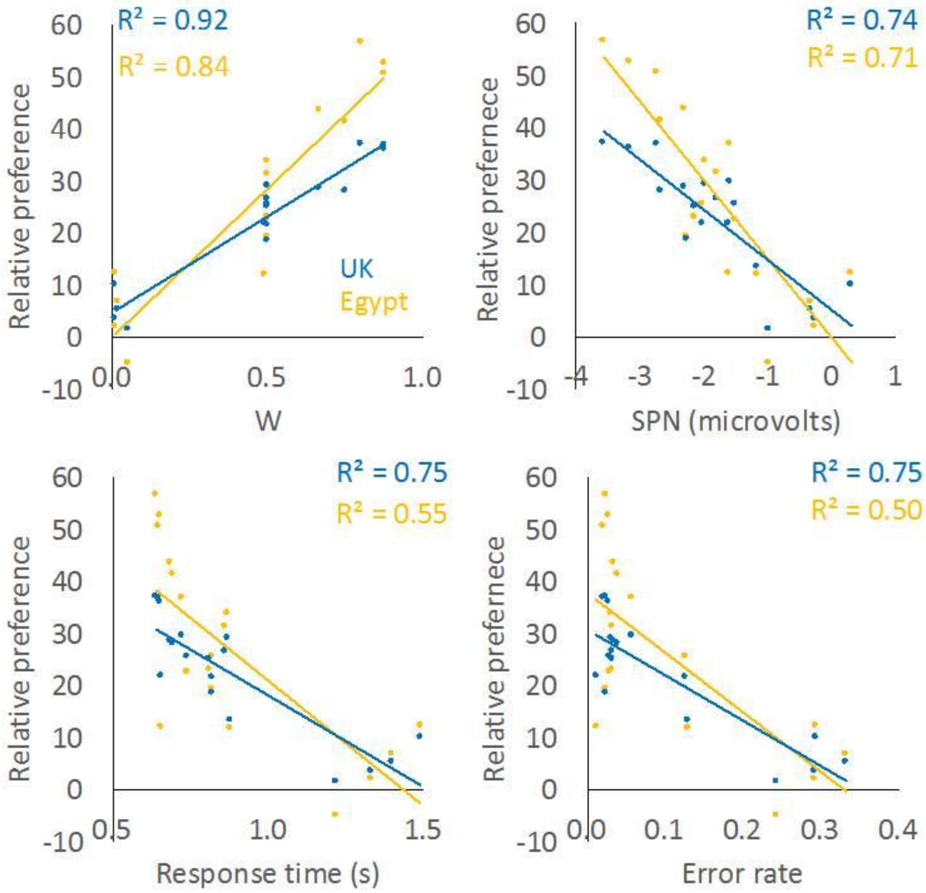
Figure 2. Relative preference correlates with four different measures of perceptual goodness. Note that a participant's *relative preference* is the difference between their mean ratings for regular patterns and mean ratings for random patterns. Positive values indicate a relative preference for regularity. Each data point represents the grand average from one regular condition and sub-experiment. The variables on the X axis were obtained using the same types of patterns, but from different participants (Makin et al., 2016). A) W-load from the holographic model. B) Grand-average SPN amplitude. C) Grand-average response time D) Grand-average Error rate. Note that an extremely negative SPN = a large neural response to symmetry, and lower response time and error rate = more efficient performance. Blue

data is from the UK sample, Yellow is from the Egyptian sample.

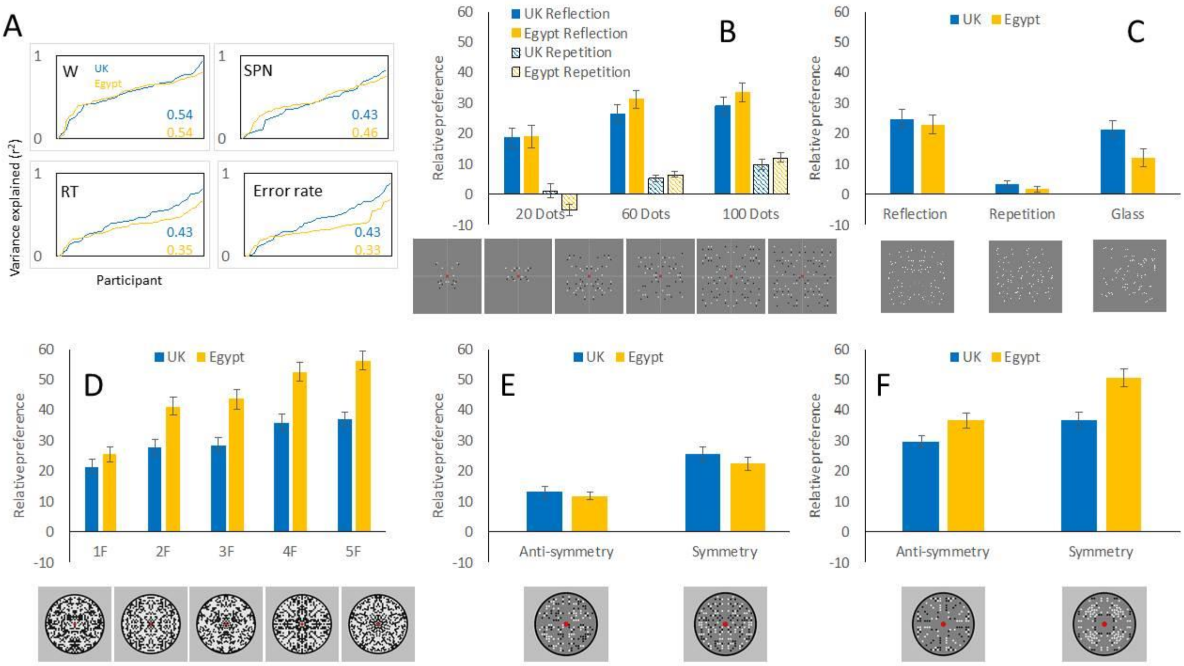
Figure 3. Results of Experiment 1. A) Analysis of individual relative preference scores. Here the proportion of variance (r^2) in relative preference was obtained from each participant, and for each measure of perceptual goodness. The participants were then organized cumulatively. **B – F)** Relative preferences of the 5 sub-experiments. Example patterns are shown in the insets below. Positive values indicate that the regular patterns of this type were liked more than random (maximum possible score = 100). Blue data is from the UK sample, Yellow is from the Egyptian sample. Error bars = +/- 1 S.E.M.

Figure 4. Stimuli and results of Experiment 2. A) Example patterns from Experiment 2. These patterns are just examples. Novel examples were generated by the program on every trial. Psymm and W-loads are indicated below. **B)** Relative preference as a function of Psymm (Error bars = +/- 1 S.E.M). **C)** Correlation between W and relative preference, and r^2 values from Individual participants, organized cumulatively. **D)** Correlation between SPN amplitude and relative preference, and r^2 from Individual participants, organized cumulatively.

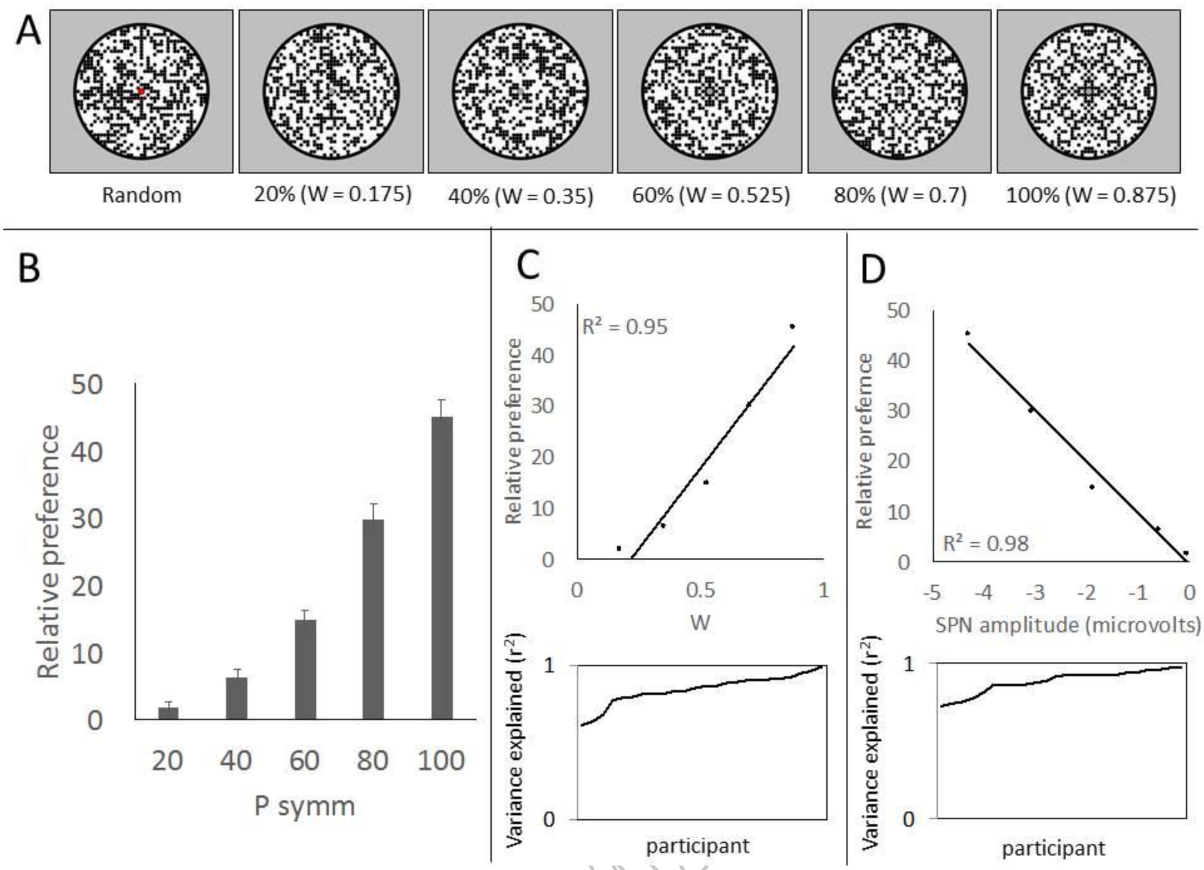

SUB EX 1
Reflection 20
W = 0.5
Reflection 60
W = 0.5
Reflection 100
W = 0.5
Repetition 20
W = 0.05
Repetition 60
W = 0.016
Reflection 100
W = 0.01
SUB EX 2
Reflection
W = 0.5
Repetition
W = 0.01
Glass
W = 0.49
SUB EX 3
1 Fold
W = 0.5
2 Fold
W = 0.75
3 Fold
W = 0.67
4 Fold
W = 0.875
5 Fold
W = 0.8
SUB EX 4
1 Fold AS
1 Fold S
W = 0.5
SUB EX5
4 Fold AS
4 Fold S
W = 0.875



R² = 0.92
R² = 0.84
UK
Egypt
Relative preference
W
R² = 0.74
R² = 0.71
Relative prefernece
SPN (microvolts)
R² = 0.75
R² = 0.55
Relative preference
Response time (s)
R² = 0.75
R² = 0.50
Relative prefernece
Error rate



A
B
C
D
E
F
W
SPN
RT
Error rate
UK
Egypt
0.54
0.54
0.43
0.46
0.43
0.35
0.43
0.33
Variance explained (r^2)
Participant
UK Reflection
Egypt Reflection
UK Repetition
Egypt Repetition
Relativepreference
20 Dots
60 Dots
100 Dots
Reflection
Repetition
Glass
1F
2F
3F
4F
5F
Anti-symmetry
Symmetry



A
Random
20% (W = 0.175)
40% (W = 0.35)
60% (W = 0.525)
80% (W = 0.7)
100% (W = 0.875)
B
Relative preference
P symm
C
R² = 0.95
Relative preference
W
Variance explained (r²)
participant
D
R² = 0.98
Relative preference
SPN amplitude (microvolts)
Variance explained (r²)
participant