

# Perceptual Similarities Among Wallpaper Group Exemplars

Peter J. Kohler<sup>1,2</sup>, Shivam Vedak<sup>3</sup>, and Rick O. Gilmore<sup>3</sup>

<sup>1</sup> York University, Department of Psychology, Toronto, ON M3J 1P3, Canada

<sup>2</sup> Centre for Vision Research, York University, Toronto, ON, M3J 1P3, Canada

<sup>3</sup>Department of Psychology, The Pennsylvania State University, Pennsylvania, USA

## Abstract

Symmetries are abundant within the visual environment, and many animal species are sensitive to visual symmetries. Wallpaper groups constitute a class of 17 regular textures that each contain a distinct combination of the four fundamental symmetries, translation, reflection, rotation and glide reflection, and together represent the complete set of possible symmetries in two-dimensional images. Wallpapers are visually compelling and elicit responses in visual brain areas that precisely capture the symmetry content of each group in humans and other primates. Here we ask to what extent *different* exemplars from the *same* wallpaper group are perceptually similar. We used an algorithm to produce a set of well-matched exemplars from 5 of the 17 wallpaper groups and instructed participants to freely sort the exemplars from each group into as many subsets as they wished based on any criteria they saw appropriate.  $P_1$ , the simplest of the 17 groups, was consistently rated more self-similar than any other group, while the other four groups, although varying in symmetry content, were comparable in self-similarity. Our results suggest that except for the most extreme case ( $P_1$ ), perceived self-similarity of wallpaper groups is not directly tied to categories of symmetry based on group theory.

key words: wallpaper groups, visual perception, behavioral sorting, self-similarity

## Introduction

Symmetry exists in an object or pattern if a transformation can be applied that maps the object/pattern onto itself. In the two-dimensional plane, the set of isometries - distance-preserving transformations, see (Liu et al., 2010) - that can give rise to symmetries are translation, reflection, rotation and glide reflection and their combinations. The *wallpaper groups* are a set of 17 regular textures, where each has a unique combination of isometries that leave the texture unchanged (Fedorov, 1891; Polya, 1924; Liu et al., 2010). Each wallpaper group therefore contains a distinct combination of four symmetry types (see Figure 1). Symmetries have been recognized as important for human visual perception since the late 19th century (Mach, 1959). Wallpaper groups are different from stimuli typically used to probe the role of symmetry in visual perception in two ways: First, they contain combinations of four symmetry types, rather than just reflection (also called mirror symmetry), which have been the focus of most vision research. Second, in wallpaper groups symmetries are repeated to tile the plane and form textures, instead of being positioned at a single image location as is usually the case with standard stimuli. These differences, and the important

fact that wallpaper groups together form the complete set of symmetries possible in the two-dimensional image plane, make wallpapers an interesting stimulus set for studying perception of visual symmetries.

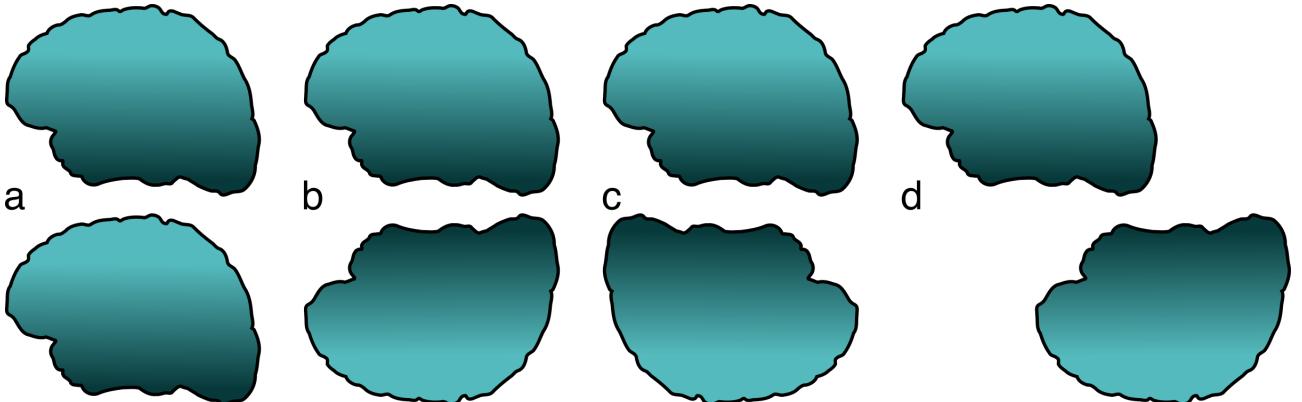


Figure 1: The four fundamental symmetry types: (a) translation, (b) reflection, (c) rotation (order 2, 180°), (d) glide reflection - translation followed by reflection over a line parallel to the direction of translation.

Brain imaging studies using functional MRI (Kohler et al., 2016) and EEG (Kohler et al., 2018; Kohler and Clarke, 2021) have shown that the human visual system carries detailed and precise representations of the symmetries within the individual wallpaper groups. Specifically, response amplitudes scale approximately linearly with the symmetry content within the wallpaper groups, across all of the possible combinations of reflection, rotation and glide reflection symmetries. Functional MRI evidence from macaque monkeys reveal similar representations in the macaque visual system, and the brain regions responding to symmetry are largely analogous between humans and monkeys, namely functionally defined regions V<sub>3</sub>, V<sub>4</sub>, VO1 and LOC (Audurier et al., 2021).

The wallpaper group representations that have been identified using brain imaging are highly complex, but do not appear to be readily available for driving conscious behaviour: Humans have limited intuitive sense of group membership for wallpaper group exemplars, as evidenced by behavioral experiments showing that although naïve observers can distinguish many of the wallpaper groups (Landwehr, 2009), they tend to sort exemplars into fewer (4-12) sets than the number of wallpaper groups, often placing exemplars from different groups into the same set (Clarke et al., 2011). Wallpaper groups are nonetheless visually compelling, and anecdotally we have observed that exemplars from a given group can be quite perceptually diverse. This observation inspired the current study. Here, we use behavioral sorting, a common technique to study perceptual categorization (Milton et al., 2008; Pothos et al., 2011), to probe the perceptual self-similarity of different exemplars from the same wallpaper group. In previous sorting experiments with wallpaper groups (e.g. (Clarke et al., 2011)) observers were shown exemplars from different wallpaper groups and their ability to correctly sort exemplars from the same group into the same subset was assessed. Our approach was different: We wanted to know the extent to which exemplars from the *same group* would be spontaneously organized into subsets, i.e. the self-similarity of exemplars from a given group. We selected five distinct wallpaper groups: *P*<sub>1</sub>, *P*<sub>3</sub>*M*<sub>1</sub>, *P*<sub>3</sub>*1**M*, *P*<sub>6</sub> and *P*<sub>6</sub>*M* (see Figure 2). All wallpaper groups consists of a lattice that is repeated to tile the plane. *P*<sub>1</sub> is the simpleste group, and contains no symmetries other than the translation generated by the repeating lattice. *P*<sub>6</sub> has rotation symmetries of order 6, 3 and 2, but no other symmetries besides translation. *P*<sub>3</sub>*M*<sub>1</sub> and *P*<sub>3</sub>*1**M* both have rotations of order 3, reflections in 3 distinct directions, and glide reflections in 3 distinct directions, but differ in terms of how these symmetries are organized in the lattice. *P*<sub>6</sub>*M* is the most complex of the groups, it has rotation symmetries of order 6, 3 and 2, reflections in 6 distinct directions, and glide reflections in 6 distinct directions. The lattice structure of the five groups is described in detail on the wallpaper group wikipedia page. The five groups selected have all been found to have high self-similarity (Clarke et al., 2011), and four of them (*P*<sub>3</sub>*M*<sub>1</sub>, *P*<sub>3</sub>*1**M*, *P*<sub>6</sub> and *P*<sub>6</sub>*M*) share the same lattice shape (see Figure 2).

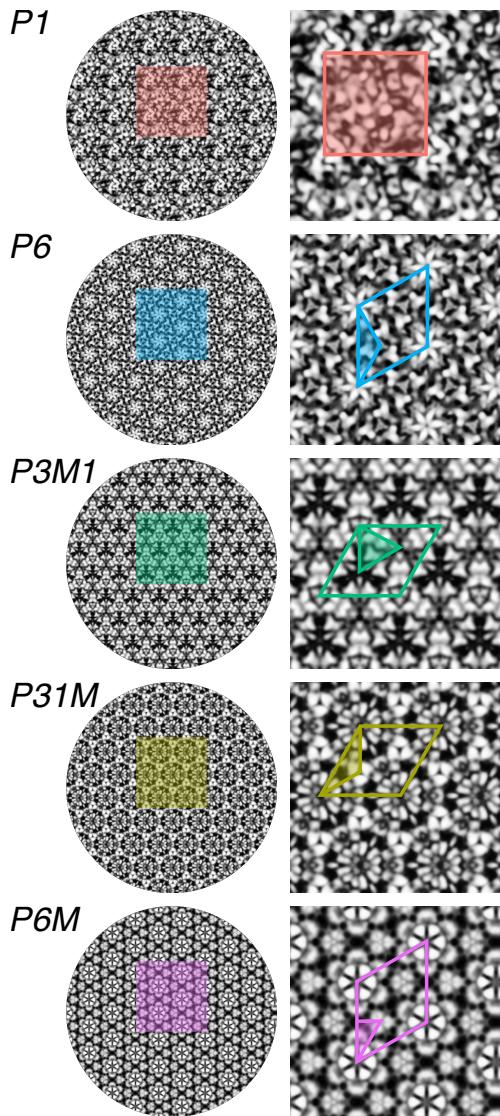


Figure 2: The fundamental region and lattice structure of the five wallpaper groups used in the study. The complete wallpaper is shown in the left-hand column with a shaded region that is repeated and enlarged in the right-hand column. The colored outline in the enlarged region indicates the repeating lattice for each group, while the shaded area indicates the fundamental region (see text). For  $P_1$  the fundamental region covers the entire lattice. Note that even though  $P_6$  and  $P_{31}M$  have the same fundamental region and lattice shapes, they differ in terms of the symmetries present within the lattice - most notably,  $P_{31}M$  contains reflection symmetry while  $P_6$  does not. The symmetry content of each group is detailed on the wallpaper group [wikipedia page](#).

Participants were given 20 exemplars, all belonging to same group (see Figure 3 for a selection of the exemplars, and the Materials and Methods section for details on how they were created) and asked to freely sort them into as many subsets as they wished. Participants sorted exemplars belonging to five different wallpaper groups, one group at a time. This approach allowed us to compare the five wallpaper groups, both in terms of how many subsets participants generated, and also in terms of the Jaccard index, a summary statistic capturing the similarity across exemplar pairs for each group. Within each group, we were also able to identify exemplar pairs that were rated as highly similar and highly dissimilar. Our main conclusion is that  $P_1$  was systematically more self-similar than the any other groups, while the other four other groups could not be distinguished on these measures. We also show that for all five groups, participants consistently group certain pairs of exemplars together, although the number of consistent pairs varies among groups. Our results open the door to further investigations into the psychological and neural mechanisms that drive perceptual similarity among wallpaper group exemplars, and indeed among

83 exemplars from different classes of structured patterns.

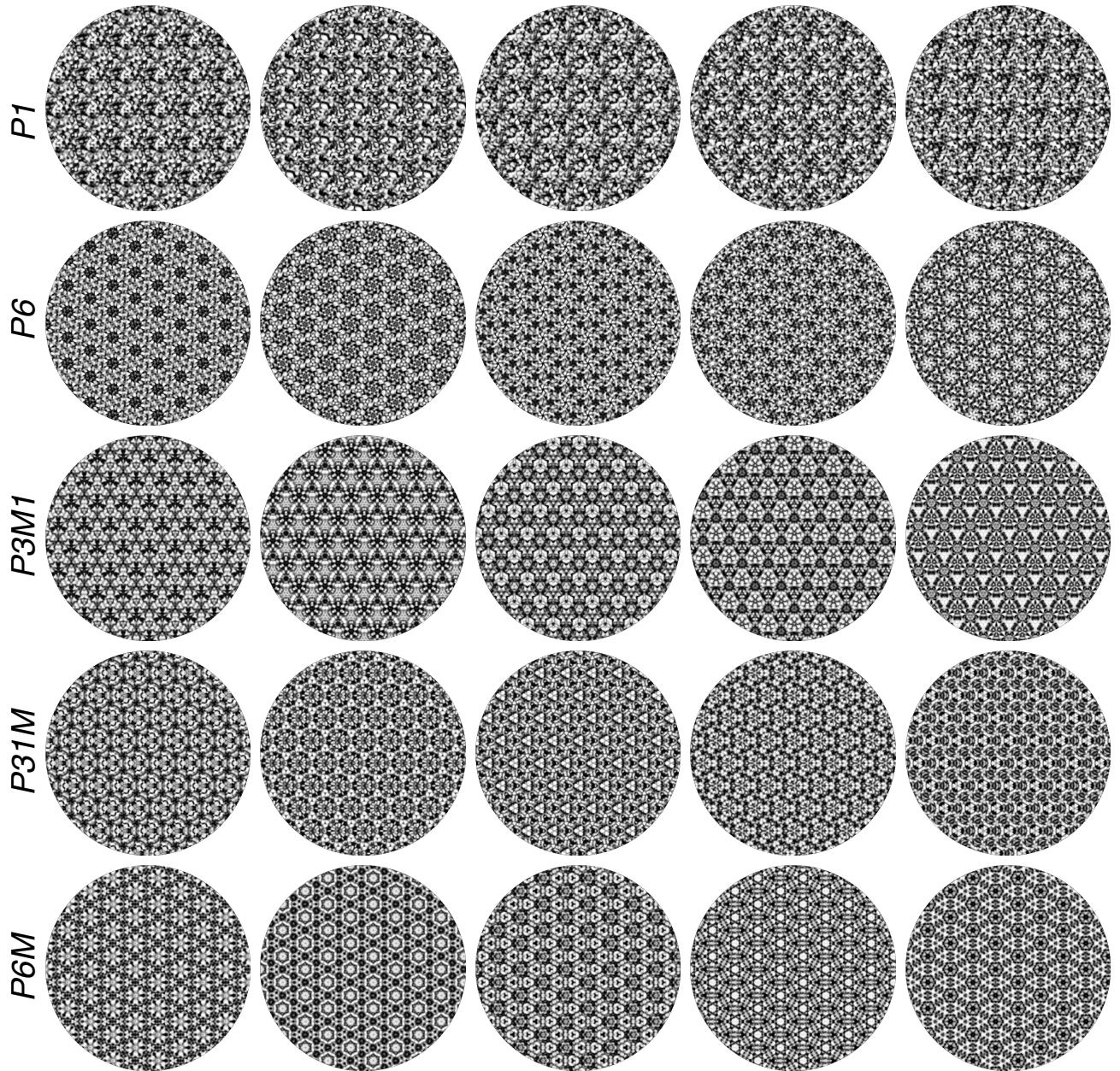


Figure 3: 5 of the 20 exemplars used for each group are shown to highlight the diversity among exemplars.

## 84 Results

85 Wallpaper group  $P_1$  was more self-similar than the other four groups. This was evident in the number of  
86 sets generated for this group across participants, which was lower for  $P_1$  (median = 3) than for the four other  
87 groups (median = 4-5, see Figure 4). We confirmed this observation statistically by running a repeated  
88 measures analysis of variance (ANOVA), which revealed a significant effect of group ( $F(4, 124) = 7.330, p <$   
89 0.0001)). Post-hoc pairwise  $t$ -tests showed that the mean number of sets was lower for  $P_1$  than all other  
90 groups ( $p_s < 0.005$ ), but no other means differed (see Table 1). Next, we computed the Jaccard index (see  
91 Materials and Methods) across participants for every pairwise combination of exemplars in each group.  
92 This provides a measure of the similarity between exemplars within each group.  $P_1$  had systematically  
93 higher Jaccard indices than the four other groups (see Figure 4), as confirmed by a repeated measures  
94 ANOVA which revealed a statistically significant effect of group ( $F(4, 495) = 20.178, p < 0.0001$ ). Post-hoc  
95  $t$ -tests showed that  $P_1$  had higher Jaccard indices than all other groups ( $p_s < 0.0001$ ; see Table 1). The

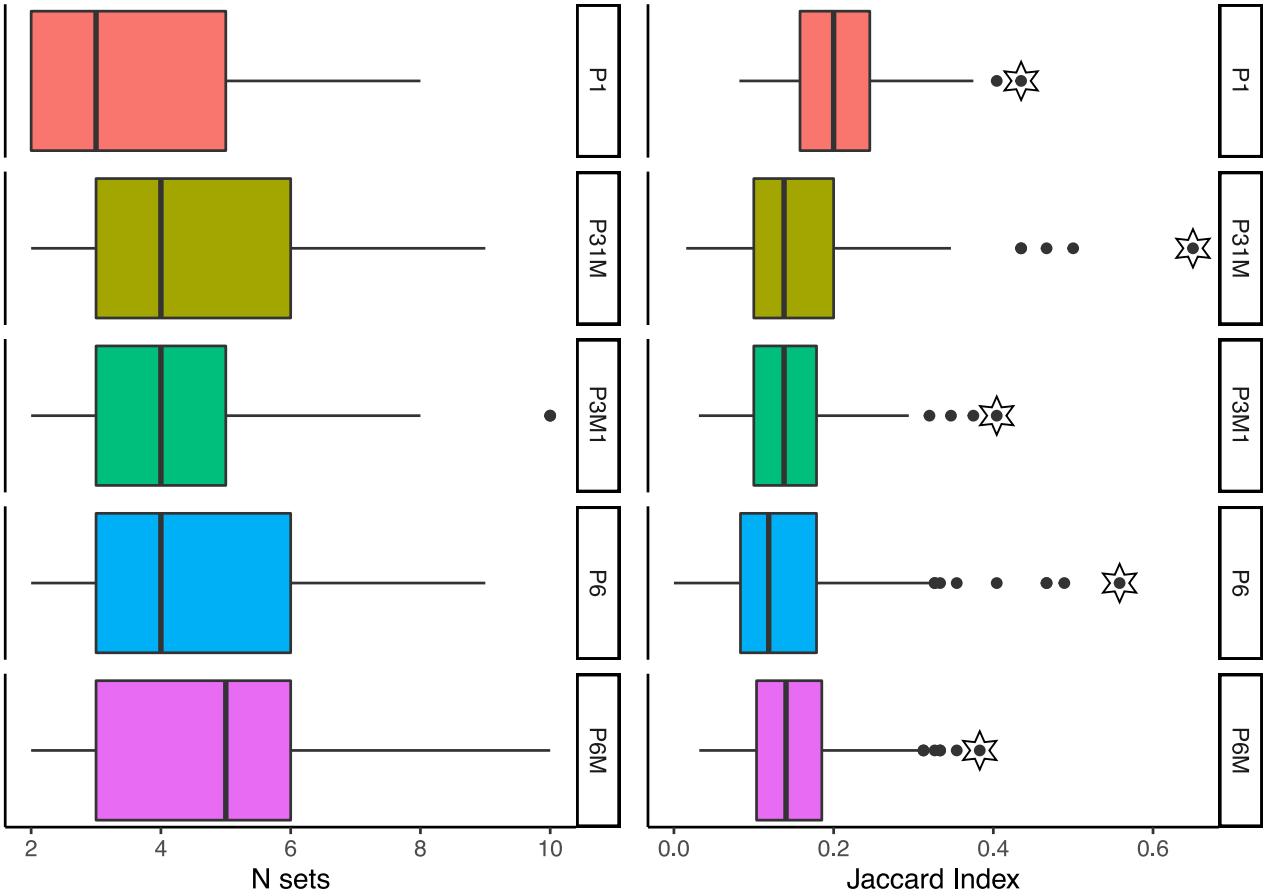


Figure 4: Left panel: Boxplots showing the number of subsets generated by participants for each of the wallpaper groups. Right panel: Boxplots showing Jaccard indices for every pairwise combination of exemplars in each of the wallpaper groups. Note that each data point here is the Jaccard index for a particular exemplar pair calculated across participants unlike the left panel where each data point is a participant. The exemplar pairs with the highest Jaccard indices have been highlighted with stars. Those outlier pairs are explored further in Figure 6. For both panels, the lower box boundary is the 25th percentile. The dark line in the box is the median. The upper box boundary is the 75th percentile. The “whiskers” show -/+ the interquartile range \* 1.5.

fact that the group ( $P_1$ ) for which fewer subsets were generated also had higher Jaccard indices than the other groups illustrates the inherent link between the two measures. For wallpaper groups where the exemplars are sorted into fewer subsets, each individual exemplar pair is more likely to be a member of the same subset, and less likely to be a member of distinct subsets. This in turn leads to higher Jaccard indices. Our pairwise  $t$ -tests also showed that  $P_{31M}$  had lower Jaccard indices than  $P_6$  ( $p = 0.037$ ). This effect does not pass our Bonferroni-corrected threshold for significance ( $\alpha < 0.005$ , but may nonetheless possibly reflect real differences in how consistently exemplars were grouped together across participants. We will explore this idea more in depth shortly. Out of the five groups tested,  $P_1$  is the only one that can be reliably differentiated based on our measures, being higher on self-similarity among the exemplars, and thus lower on diversity among exemplars.

pairs	number of sets			Jaccard Index		
	t	p	D	t	p	D
$P_1$ vs $P_{31}M$	-2.981	0.0034	-0.734	5.641	<0.0001	0.579
$P_1$ vs $P_3M_1$	-3.423	0.0008	-0.843	7.233	<0.0001	0.742
$P_1$ vs $P_6$	-4.748	<0.0001	-1.169	7.734	<0.0001	0.794
$P_1$ vs $P_6M$	-4.553	<0.0001	-1.132	6.946	<0.0001	0.713
$P_{31}M$ vs $P_3M_1$	-0.442	0.6595	-0.109	1.592	0.1117	0.163
$P_{31}M$ vs $P_6$	-1.767	0.0797	-0.435	2.094	0.0366	0.215
$P_{31}M$ vs $P_6M$	-1.600	0.1120	-0.398	1.305	0.1921	0.134
$P_3M_1$ vs $P_6$	-1.325	0.1875	-0.326	0.502	0.6160	0.051
$P_3M_1$ vs $P_6M$	-1.163	0.2470	0.289	-0.287	0.7745	-0.029
$P_6$ vs $P_6M$	0.150	0.8814	0.037	-0.788	0.4307	-0.081

Table 1: Results of post-hoc pairwise  $t$ -tests on number of sets and Jaccard Indices. Degrees-of-freedom was 945 for the Jaccard Index test. For the number of sets test, degrees-of-freedom had to be adjusted to account for the fact that one participant did not report number of sets for P6M (see Materials and Methods), and ranged between 127.0 and 127.1.

consistent pairings		
group	$p < 0.01$	$p < 0.0001$
$P_1$	6	1
$P_{31}M$	17	10
$P_3M_1$	12	3
$P_6$	17	11
$P_6M$	15	4

Table 2: Number of consistent pairings at two different  $\alpha$ -levels for the five groups.

In order to quantify the extent to which exemplars were consistently grouped together, we ran a permutation analysis in which exemplar labels were shuffled among the sets generated for each participant (see Materials and Methods). This provides, for each group, the expected distribution of Jaccard indices for every pairwise combination of exemplars, if exemplars were assigned randomly to subsets. And the analysis allows us to compute an empirical  $z$ -score that expressed the extent to which a given pair of exemplars deviates from random assignment.

Because the random distribution is generated by shuffling exemplars across the specific sets generated by each participant for each group, this  $z$ -score is independent of the number of sets. If for a given group, none of the pairs deviate significantly from the random distribution, it would indicate that no exemplar pairs were consistently grouped together across participants. To estimate the extent to which this is the case, we look at the distribution of  $z$ -scores across the pairs for each group, as plotted in Figure 5, and count the number of pairs for each group for which the  $p$ -value associated with the threshold exceeds a given  $\alpha$  value. At a threshold of  $\alpha = 0.01$ , several pairs survive for all groups, and even at a much more conservative criterion of  $\alpha = 0.0001$  most groups have more than one pairing that survives (see Table 2). It is worth noting that the latter threshold ( $\alpha = 0.0001$ ) is lower than the  $\alpha$  associated with a Bonferroni correction within group, given that there are 190 pairs per group:

$$\alpha = \frac{0.05}{190} = 0.0003$$

122 So we conclude that for several exemplar pairs, participants are consistent in how they tend to pair the  
 123 exemplars. It is interesting to consider that this measure of consistency might provide another way of  
 124 differentiating wallpaper groups in terms of perceptual self-similarity. While groups  $P_{31}M$ ,  $P_3M_1$ ,  $P_6$  and  
 125  $P_{6M}$  have comparable Jaccard scores (see Figure 4), they differ in the number of consistent pairings, with  
 126  $P_{31}M$  and  $P_6$  producing more consistent pairs than the other two (see Figure 5).

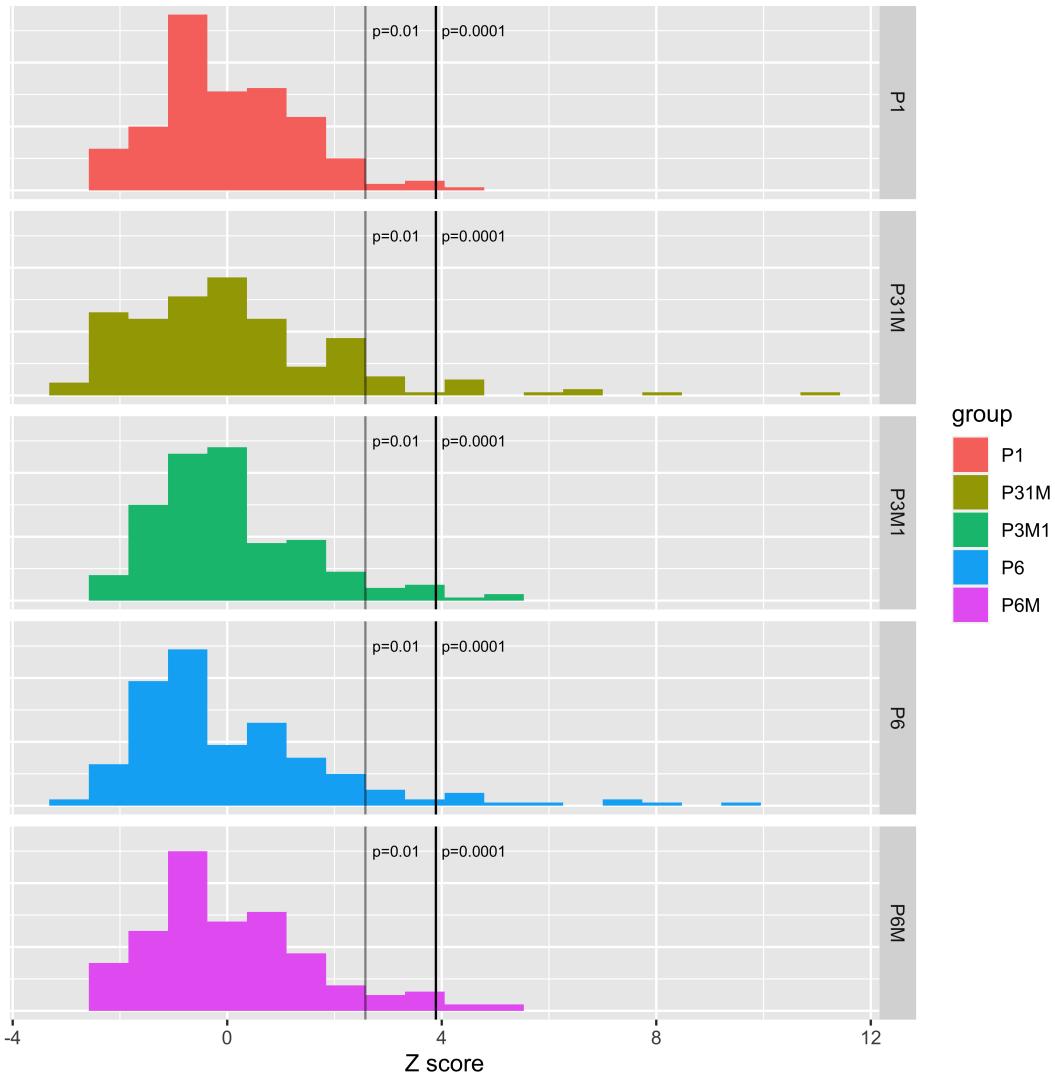


Figure 5: Distribution of  $z$ -scores across the 190 pairs in each of the five wallpaper groups. The two lines indicate the  $z$ -scores associated with  $\alpha$  of 0.01 and 0.0001, respectively.

127 The Jaccard indices also allow us to focus on exemplar pairs that have a high level of similarity relative  
 128 to the rest of the pairs in the set. We do this by identifying outliers pairs from each group in term of Jaccard  
 129 indices, as identified with stars in Figure 4. Because the Jaccard indices are computed across participants,  
 130 these outliers are also among the pairs most consistently sorted together, as identified in Figure 5. For each  
 131 exemplar in each outlier pair, we can visualize the pairwise similarity (as measured by the Jaccard index)  
 132 to every other exemplar in the set (see Figure 6). That is, we can visualize portions of the network of  
 133 perceived similarity within a set of exemplars. Future work could probe the extent to which networks of  
 134 perceived similarity have similar structure across wallpaper groups and examine what perceptual features  
 135 best account for participants' perceptions of exemplar similarity.

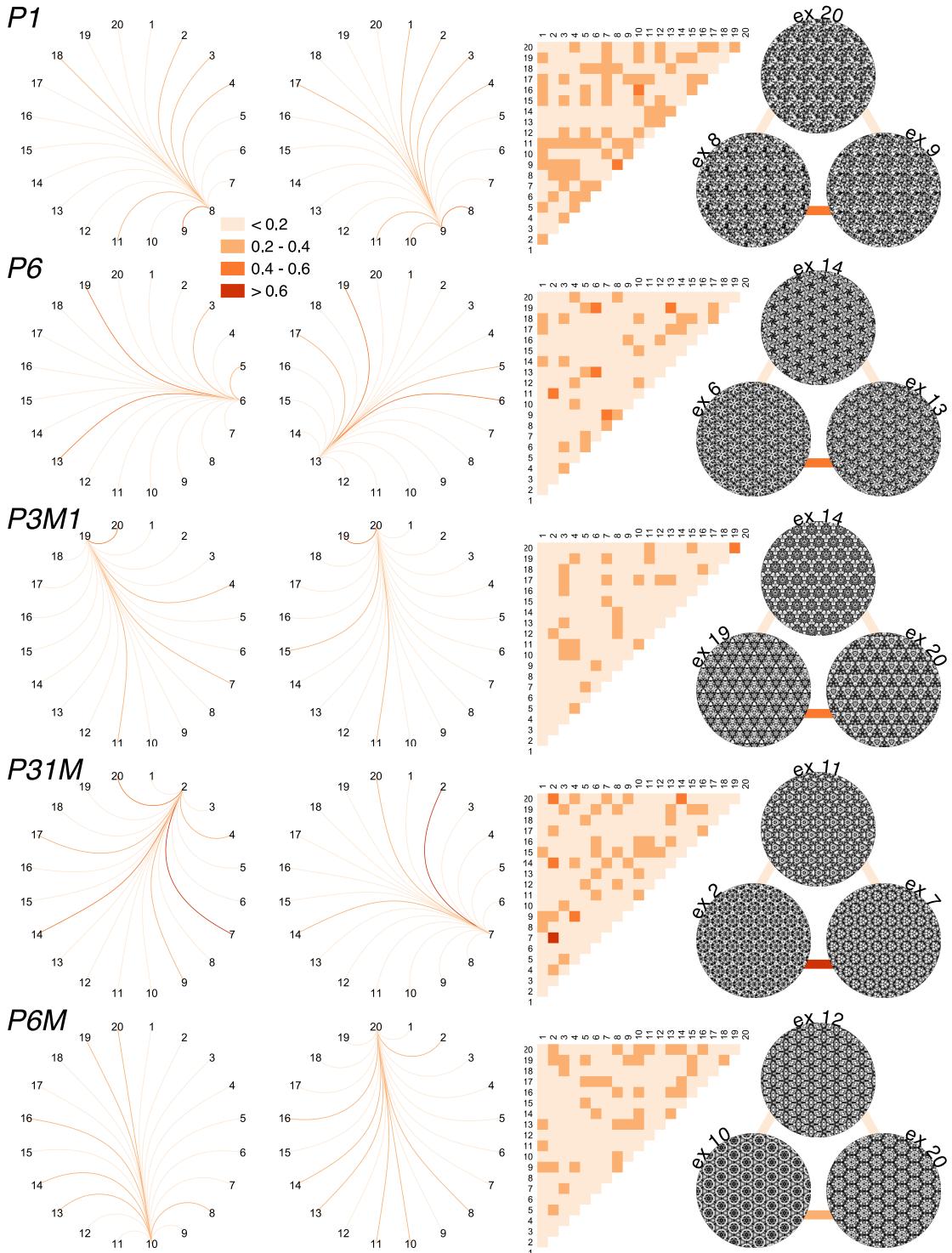


Figure 6: For each wallpaper group, we identified the two most self-similar exemplars, the same pair that is indicated by the right-most datapoint for each group in Figure 5. The two circular network plots are showing the pairwise similarities between those two exemplars and every other exemplar in the set. The pairwise similarities across all exemplars are plotted as a similarity matrix and on the rightmost side of the plot, the two most self-similar exemplars (bottom) are plotted with the exemplar that was least similar to both (top). The connecting lines between the exemplars indicate the similarity.

## 136 Discussion

137 Previous work has demonstrated that visual cortex of both humans and macaque monkeys carries highly  
 138 detailed representations of the symmetries within wallpaper groups, as evidenced by systematic differences

139 in the magnitude of the response elicited by different groups (Kohler et al., 2016; Kohler and Clarke,  
140 2021; Audurier et al., 2021). This distinction between groups can also be observed in psychophysical  
141 threshold measurements (Kohler and Clarke, 2021), although observers may not have a strong awareness  
142 of the wallpaper group membership of individual exemplars (Clarke et al., 2011). In the current study, we  
143 explored a new piece of the story about how wallpaper groups are processed by the visual system, namely the  
144 issue of how self-similar different exemplars from the *same* wallpaper group appear to untrained observers.  
145 We tested this by asking participants to spontaneously sort 20 exemplars from each of five wallpaper groups  
146 into different subsets.

147 Our first finding concerns the number of subsets generated for each group. We find that  $P_1$  is divided  
148 into fewer subsets than the other four groups. This indicates that the limited complexity of this group,  
149 which contains only translation symmetry, has a direct effect of the number of distinct subsets. The rela-  
150 tionship between complexity / symmetry content and number of subsets produced is not straightforward,  
151 however, as indicated by the fact that  $P6M$  is not consistently grouped into more subsets than  $P6$ ,  $P3M_1$  and  
152  $P31M$ , despite the fact that these other groups all contain fewer symmetries than  $P6M$ . We speculate that  
153 this lack of further differentiation is a result of an upper limit on how additional complexity can influence  
154 perceptual self-similarity. However, future work with additional wallpaper groups, including groups that  
155 are relatively low on complexity by high on self-similarity (e.g.,  $P_2$  and  $PMM$ , see (Clarke et al., 2011)), is  
156 needed to draw firm conclusions about this hypothesis.

157 It is important to note that  $P6$ ,  $P3M_1$  and  $P31M$  all consistently generate weaker brain activity than  $P6M$ ,  
158 and produce higher thresholds in a symmetry detection task (Kohler and Clarke, 2021). Our results would  
159 therefore suggest that there is no clear relationship between the strength of the visual system's response to  
160 symmetries in wallpaper group, and the perceptual self-similarity of each individual group. Future work  
161 should explore this more closely, and look for neural correlates of similarity among exemplars from the  
162 same group.

163 We also computed Jaccard indices that, for every possible exemplar pair, expresses the frequency of  
164 those two exemplars being grouped together. As described above, the average Jaccard index for a group  
165 is inherently linked to the number of subsets produced for that group, because fewer subsets mean that  
166 exemplars are more likely to be made members of the same pair, and less likely to be made members of  
167 the same pair. It is therefore not surprising that we find the same general pattern for Jaccard indices and  
168 number of subsets, namely that  $P_1$  has higher indices than the other groups. The advantage of the Jaccard  
169 indices, however, is that they allow us to conduct a permutation analysis that quantifies the extent to which  
170 pairs of exemplars are consistently grouped together across participants, independent of the number of  
171 sets produced for a given group. It is important to note that consistency in the choice of which exemplars  
172 to group into subsets is not an unavoidable consequence of our experimental design, and it does not follow  
173 naturally from the results described so far. It would be perfectly possible for participants to group the sets  
174 together, producing fewer subsets for  $P_1$  as observed, but exhibit no consistency across participants at all.  
175 That is not what we see, however. Even when setting a conservative threshold, all five groups produce  
176 one or more pairs that are consistently grouped together, demonstrating that the sorting of exemplars into  
177 subsets is not done randomly or arbitrarily across the participants. Rather, different individuals agree to  
178 some extent on which exemplars belong together. Because our measure of consistency is independent of  
179 the number of subsets produced for a given group, it allows us to show that although  $P_1$  has the highest  
180 overall Jaccard indices (as a result of the fewer sets produced for this group), it in fact produces fewer  
181 consistent pairs than other groups (see Table 2). Indeed, participants made few comments about their own  
182 sorting strategies, but most observed that  $P_1$  exemplars were the most difficult to sort because of the lack  
183 of readily apparent features that were consistent across exemplars.

184 In sum, we find consistencies in the way that untrained human observers sort wallpaper images. Ob-  
185 servers sort exemplars with translational symmetry alone ( $P_1$ ) into smaller numbers of sets than exemplars

186 with rotation or reflection symmetry. On average, pairs of  $P_1$  exemplars are sorted together more often  
187 than exemplars from other wallpaper groups. At the same time, some specific exemplar pairs from wall-  
188 paper groups with 3- or 6-fold rotational or reflection symmetry are sorted together substantially more  
189 often than predicted by chance.

190 We note that the spontaneous sorting task our observers engaged in has less intrinsic structure  
191 than some other tasks used to study similar questions like oddball detection (Landwehr, 2009; Hebart et al.,  
192 2020; Landwehr, 2011), and thus may involve somewhat different perceptual and cognitive processes. In  
193 particular, wallpaper group exemplars have a reduced dimensionality relative to natural objects. Even so,  
194 large scale evaluations of how human observers perceive similarity in natural objects yield dimensions that  
195 appear to relate to the strict regularities observed in wallpapers: round shape, patterning, and repetition  
196 (Hebart et al., 2020). In future work, it would be interesting to explore whether different behavioral tasks  
197 yield comparable similarity spaces, or more generally, how task demands shape similarity judgments.

198 In conclusion, our results suggest that human observers show sensitivity to the dimensions of 2D  
199 symmetry (translation, rotation, and reflection) embedded in wallpaper exemplars. However, their sorting  
200 behavior shows only weak evidence that group-theoretic measures of symmetry influence the perception  
201 of self-similarity. These results contribute to a small, but growing literature on the perception of visual  
202 aesthetics (Carneiro et al., 2012; Graham et al., 2010; Friedenberg, 2012; Laine-Hernandez and Westman,  
203 2008; Richards, 1972) where symmetry is one of many contributing factors.

## 204 Materials and Methods

### 205 Participants

206 33 participants (9 Male, 24 Female), ranging in age between 18 and 35 completed this study. Participants  
207 were recruited from the undergraduate participant pool in the Department of Psychology at the  
208 The Pennsylvania State University. All participants had self-reported 20/20 or corrected to 20/20 vision.  
209 We obtained written consent to participate from all participants under procedures approved by the In-  
210 stitutional Review Board of The Pennsylvania State University (#38536). The research was conducted  
211 according to the principles expressed in the Declaration of Helsinki. Participants include  $n=11$  collected  
212 and described in (Vedak, 2014), plus an additional group collected at a later date using the same protocol.

### 213 Stimulus Generation

214 Five wallpaper groups ( $P_1$ ,  $P_6$ ,  $P_{31}M_1$ ,  $P_{31}M$ , and  $P6M$ ) were selected for use in the study. The selection was  
215 motivated partially by the fact that all five groups have previously been shown to be high in self-similarity  
216 (Clarke et al., 2011), and partially by the fact that  $P_{31}M_1$ ,  $P_{31}M$ ,  $P_6$  and  $P6M$  all share the same lattice shape. We  
217 also found it interesting that while  $P_6$ ,  $P_{31}M_1$  and  $P_{31}M$  differ in their symmetry content, all are subgroups  
218 of  $P6M$  with index 2, which means that  $P6M$  can be generated by adding one additional transformation to  
219  $P_6$ ,  $P_{31}M_1$  or  $P_{31}M$  (Kohler and Clarke, 2021). 20 exemplars from each of these five wallpaper groups were  
220 generated using a modified version of the methodology developed by Clarke and colleagues (Clarke et al.,  
221 2011) that we have described in detail elsewhere (Kohler et al., 2016). Briefly, exemplars belonging to each  
222 group were generated by starting with a random-noise patch, which was then repeated and transformed  
223 to tile the image plane, in accordance with the symmetry axes and geometric lattice specific to each group.  
224 The use of noise patches as the starting point for stimulus generation makes it possible to create an almost  
225 unlimited number of distinct exemplars from each wallpaper group. To make individual exemplars as  
226 similar as possible we replaced the power spectrum of each exemplar with the median across exemplars  
227 within a group. These images were printed onto white cardstock and cut into squares, allowing participants

228 to manipulate the orientation of the images during the sorting tasks. Five exemplars from each group are  
229 shown (in reduced size) in Figure 3.

## 230 Procedure

231 Participants were presented with the 20 exemplars of a single wallpaper group (i.e. P1, P<sub>3</sub>M1, P<sub>31</sub>M, P6,  
232 P6M) and instructed to sort them into subsets by placing them into piles. Participants were advised to  
233 sort the exemplars into as many piles as they deemed necessary based on whatever criteria they desired.  
234 There were no time constraints placed on this sorting task, and the participants were allowed to move  
235 exemplars between piles until they were satisfied with their classification. This method was then repeated  
236 for the remaining four wallpaper groups for each participant, with group presentation order randomized  
237 between participants. These tasks were carried out on a large table with sufficient space to randomly lay  
238 out all twenty exemplars of each set, illuminated by normal overhead room lighting. Upon completion  
239 of each sorting task, participants were asked to verbalize which features they used to sort the exemplars.  
240 After completion of all five sorting tasks, participants were asked which if they had a distinct method for  
241 sorting the images, and if any wallpaper group was particular easy or difficult to sort.

## 242 Generating the Jaccard Index

243 The data was prepared for analysis by creating one binary variable for each subset created by each par-  
244 ticipant within a sorting task. Then, each exemplar was assigned a value of one (1) if it was included in a  
245 subset, or a value zero (0) if it was not. Next, the similarity of each pair of exemplars within a sorting task  
246 was calculated using the Jaccard index, a measure of similarity and diversity for binary data. This index  
247 is calculated by the equation

$$J = \frac{x}{x + y + z}$$

248 with x representing the number of subsets that contained both exemplars, and y and z the number of  
249 subsets that contain only one exemplar of the pair ([Capra, 2005](#)), across participants. Thus, the Jaccard  
250 index is the ratio of the number of subsets containing both exemplars of a pair to the number of subsets  
251 containing at least one of the exemplars of a pair, thereby excluding subsets with joint absences.

## 252 Statistical Analysis

253 We tested for differences between the five wallpaper groups tested in terms of number of sets produced  
254 and Jaccard indices, by running repeated measures analyses of variance (rmANOVA) with group as a fixed  
255 factor and participant as a random factor. We then tested the extent to which differences between specific  
256 pairs of wallpaper groups contributed to any rmANOVA effects found, by running post-hoc paired *t*-tests  
257 comparing every possible pairing of the wallpaper groups, for both number of sets and Jaccard Indices.  
258 Because there were 10 possible pairings of the groups, we applied Bonferroni-correction and adjusted our  
259  $\alpha$ -level so that each *t*-test was only considered significant if  $p < 0.005$ . Number of sets data was missng for  
260 one participant for wallpaper group P6M, so the degrees of freedom used in the paired *t*-tests was adjusted  
261 using the Kenward-Roger technique—the default in the *emmeans* package in R.

262 We ran a permutation analysis in order to quantify the extent to which pairs of exemplars were consis-  
263 tently grouped together, across participants. This involved generating a randomized dataset, as follows.  
264 For each participant and wallpaper group, we randomized which specific exemplars were sorted together.  
265 This retained the basic structure of each participants' sorting data—the number of subsets created—but  
266 randomized the relationship between specific wallpaper exemplars that were sorted together across the  
267 participants. We then created 1,000 such permuted datasets, and calculated the Jaccard index for each  
268 exemplar pair within each group for each of the permuted datasets. This permitted the calculation of

269 an empirical Jaccard index based on the permuted data from which distributional statistics like  $z$  could be  
270 calculated. The observed Jaccard indices for each exemplar pair were then compared to the empirically-  
271 derived reference distribution to determine which exemplar pairs were sorted together more frequently  
272 than chance would predict.

## 273 References

- 274 Audurier, P., Héjja-Bertrand, Y., De Castro, V.  
275 Kohler, P. J., Norcia, A. M., Durand, J.-B., and Cottetereau, B. R. (2021). Symmetry Processing in the  
276 Macaque Visual Cortex. *Cerebral Cortex*, (bhab3583).  
277
- 278 Capra, M. G. (2005). Factor Analysis of Card Sort Data: An Alternative to Hierarchical Cluster Analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 49(5):691–695.  
279 Publisher: SAGE Publications Inc.
- 280 Carneiro, G., da Silva, N. P., Del Bue, A., and Costeira, J. P. (2012). Artistic image classification:  
281 An analysis on the PRINTART database. In *Computer Vision – ECCV 2012*, pages 143–157. Springer  
282 Berlin Heidelberg.
- 283 Clarke, A. D. F., Green, P. R., Halley, F., and Chantler, M. J. (2011). Similar symmetries: The  
284 role of wallpaper groups in perceptual texture similarity. *Symmetry*, 3(4):246–264.
- 285 Fedorov, E. (1891). Symmetry in the plane. In *Zadaniya po piski Imperatorskogo S. Peterburgskogo Mineralogicheskogo Obshchestva [Proc. S. Peterb. Mineral. Soc.]*, volume 322, pages 345–390.
- 286 Friedenberg, J. (2012). Aesthetic judgment of triangular shape: compactness and not the golden ratio determines perceived attractiveness. *i-Perception*, 3(3):163–175.
- 287 Graham, D. J., Friedenberg, J. D., Rockmore, D. N., and Field, D. J. (2010). Mapping the similarity space of paintings: Image statistics and visual perception. *Visual cognition*, 18(4):559–573.
- 288 Hebart, M. N., Zheng, C. Y., Pereira, F., and Baker, C. I. (2020). Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nature human behaviour*, 4(11):1173–1185.
- 289 Kohler, P. J., Clarke, A., Yakovleva, A., Liu, Y., and Norcia, A. M. (2016). Representation of maximally regular textures in human visual cortex. *The Journal of Neuroscience*, 36(3):714–729.
- 290 Kohler, P. J. and Clarke, A. D. F. (2021). The human visual system preserves the hierarchy of two-dimensional pattern regularity. *Proceedings of the Royal Society B: Biological Sciences*, 288(1955):20211142. Publisher: Royal Society.
- 291 Kohler, P. J., Cottetereau, B. R., and Norcia, A. M. (2018). Dynamics of perceptual decisions about symmetry in visual cortex. *NeuroImage*, 167(Supplement C):316–330.
- 292 Laine-Hernandez, M. and Westman, S. (2008). Multifaceted image similarity criteria as revealed by sorting tasks. *Proceedings of the American Society for Information Science and Technology*, 45(1):1–14.
- 293 Landwehr, K. (2009). Camouflaged symmetry. *Perception*, 38:1712–1720.
- 294 Landwehr, K. (2011). Visual discrimination of the plane symmetry groups. *Symmetry*, 3(2):207–219.
- 295 Liu, Y., Hel-Or, H., Kaplan, C. S., and Van Gool, L. (2010). Computational symmetry in computer vision and computer graphics. *Foundations and Trends® in Computer Graphics and Vision*, 5(1–2):1–195.
- 296 Mach, E. (1959). *The Analysis of Sensations* (1897). English transl., Dover, New York.
- 297 Milton, F., Longmore, C. A., and Wills, A. J. (2008). Processes of overall similarity sorting in free classification. *Journal of experimental psychology. Human perception and performance*, 34(3):676–692.
- 298 Polya, G. (1924). Xii. Über die analogie der kristallsymmetrie in der ebene. *Zeitschrift für Kristallographie-Crystalline Materials*, 60(1):278–282.
- 299 Pothos, E. M., Perlman, A., Bailey, T. M., Kurtz, K., Edwards, D. J., Hines, P., and McDonnell, J. V.

- 346 (2011). Measuring category intuitiveness in unconstrained  
347 categorized tasks. *Cognition*, 121(1):83–100.
- 348
- 349 Richards, L. G. (1972). A multidimensional scaling analysis of judged similarity of complex forms
- 350
- 353 from two task situations. *Perception & psychophysics*, 12(2):154–160.
- Vedak, S. (2014). The salience of lower-order features in highly self-similar wallpaper groups. *Honors thesis, The Pennsylvania State University*.