Acta Psychologica 174 (2017) 48–53

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

A preference for some types of complexity comment on "perceived beauty of random texture patterns: A preference for complexity"

Nicolas Gauvrit [a,*], Fernando Soler-Toscano [b], Alessandro Guida [c]

[a] Human and artificial Cognition Lab, EPHE, Paris, France
[b] University of Seville, Spain
[c] LP3C, University of Rennes 2, France

ARTICLE INFO

Article history:
Received 6 October 2016
Received in revised form 19 January 2017
Accepted 23 January 2017
Available online 2 February 2017

Keywords:
Complexity
Aesthetics
Random patterns
Entropy
Algorithmic information theory
Algorithmic complexity

ABSTRACT

In two experiments, Friedenberg and Liby (2016) studied how a diversity of complexity estimates such as density, number of blocks, GIF compression rate and edge length impact the perception of beauty of semi-random two-dimensional patterns. They concluded that aesthetics ratings are positively linked with GIF compression metrics and edge length, but not with the number of blocks. They also found an inverse U-shaped link between aesthetic judgments and density. These mixed results originate in the variety of metrics used to estimate what is loosely called "complexity" in psychology and indeed refers to conflicting notions. Here, we reanalyze their data adding two more conventional and normative mathematical measures of complexity: entropy and algorithmic complexity. We show that their results can be interpreted as an aesthetic preference for low redundancy, balanced patterns and "crooked" figures, but not for high algorithmic complexity. We conclude that participants tend to have a preference for some types of complexity, but not for all. These findings may help understand divergent results in the study of perceived beauty and complexity, and illustrate the need to specify the notion of complexity used in psychology. The field would certainly benefit from a precise taxonomy of complexity measures.

1. Introduction

Within the long-standing line of research investigating the human judgment of beauty (Berlyne, 1971; Birkhoff, 1932; Eysenck, 1940), complexity has been a prominent issue with contradictory conclusions (for a brief review see Forsythe, Nadal, Sheehy, Cela-Conde & Sawey, 2011). We believe that two factors contribute to this heterogeneity (for a similar view see Nadal, Munar, Marty & Cela-Conde, 2010). Firstly, many studies on the perception of beauty have used pictures, paintings, portraits or natural world objects to increase ecological validity (e.g., Krupinski & Locher, 1988; Messinger, 1998; Nicki, Lee, & Moss, 1981; Osborne & Farley, 1970). By doing so, they have introduced unwanted confounding factors. Studies based on more abstract stimuli are probably easier to interpret. In this respect, experiments based on non figurative material (Aitken, 1974; Ichikawa, 1985; Markovic & Gvozdenovic, 2001), such as two-dimensional binary grids (i.e., grids with black and white cells, e.g., Palumbo, Ogden, Makin, & Bertamini, 2014; Spehar, Clifford, Newell, & Taylor, 2003; Bertamini, Makin, & Pecchinenda, 2013), are of special interest as they are easily modeled in a mathematical sound way as binary matrices, which gives access to well-defined measures of complexity. Secondly, what is exactly meant by "complexity" is variable from one study to another, as many definitions of complexity exist. This is problematic as different types of complexity might result in different outcomes.

In a recent paper, Friedenberg and Liby (2016) investigated how participants rate the beauty of artificial 2-dimensional binary grids for various levels of complexity. The authors used different indices of complexity, listed below:

- The *number of parts* or blocks in the pattern (a block is a maximal subset of adjacent black cells; see Appendix for examples).
- The *total edge length,* which is the perimeter of the figure defined by the black cells (see Appendix for examples).
- The *GIF compression metrics*, defined as GIF/BMP where "GIF" is the size of the image file in GIF format and "BMP" the size of the BMP image file of a given grid. BMP is a non-compressed format, and GIF is a lossless compressed format based on the Lempel-Ziv-Welch

* Corresponding author.
E-mail address: ngauvrit@me.com (N. Gauvrit).

http://dx.doi.org/10.1016/j.actpsy.2017.01.007
0001-6918/

(LZW) compression algorithm (Donderi, 2006; Forsythe, Sheehy, & Sawey, 2003). Thus, the GIF compression metrics measures the non-redundancy of sub-patterns.
- The *density*, defined as the proportion of black cells in the grid.

Results showed that participants had a preference for evenly balanced patterns with large total edge length and high GIF compression metrics.

Crucially, in the introduction Friedenberg and Liby (2016) contrasted two fundamental notions of complexity, that is, entropy (to wit, how unbalanced a pattern is), and algorithmic complexity (to wit, how compressible a pattern is), but did not use them subsequently. Here, we reanalyze the data of both experiments using these two more conventional measures of complexity. We show that the general notion that people have a preference for complexity must be specified: they do show a preference for high entropy and more crooked figures, but there is no evidence that they prefer high algorithmic complexity per se. These results may help to understand why contradictory results have been found in previous research (Forsythe et al., 2011; Nadal et al., 2010), and call for a cautious definition of complexity metrics.

2. General method

2.1. Procedure

In their two experiments, Friedenberg and Liby (2016) used a series of square grids filled with white and black cells. Participants were asked to consider the white cells as the background, and the black cells as the figure. The grids (patterns) were pseudo-random, with densities fixed at 10, 20, 30, …, 100%. In Experiment 1, five 10×10 grids of each density (total: 50 grids) were used. In Experiment 2, ten 15×15 grids of each density (total: 100 grids) were used.

2.2. Measures

2.2.1. Subjective beauty

Participants were asked to rate each grid in terms of their subjective beauty. The mean rating of each grid was used as a measure of subjective beauty. The 4 above-mentioned metrics of complexity (density, number of parts, edge length and GIF compression) were also used in our re-analysis. Below we describe them and how they relate to (Shannon information) entropy and (algorithmic) complexity.

2.2.2. Edge length and number of parts

As the definition of edge length was not fully detailed in the original paper by Friedenberg and Liby (2016), we re-computed it with our own algorithm using R (Version 2.14; CRAN project; R Development Core Team, 2013). We also re-computed the number of parts (see Appendix for the R-codes). Previous studies have demonstrated a preference for shapes with larger edge length with different material: polygons (Friedenberg & Bertamini, 2015).

The number of parts is a natural measure of "complexity". The more parts in the pattern, the more complex it might look on average. Nevertheless, it would be more accurately described as a measure of scatter, that is, how evenly spread the black cells are. For instance, a chessboard pattern achieves maximum possible number of parts, whereas it is usually not considered complex.[1]

Edge length may look similar to the number of parts, but a single part with a given number of black cells can actually have very different edge lengths, depending on how crooked it is. Regular parts such as squares have small edge length, while more chaotic parts may have very large edge length (see Appendix, Fig. 1). Edge length could be described as a measure of how crooked a figure is, which is of course linked to its algorithmic complexity as well as its entropy.

2.2.3. Entropy

The term "entropy" has been used in various ways in the literature. In some cases (for instance, with material difficult to describe in a mathematical simple form), it is used with the generic sense of "disorder" (e.g., Gillam & Grove, 2011). In other cases, especially when binary sequences or grids are used, "entropy" refers either to the classical Shannon (first order) information entropy (e.g., Hirsh, Mar, & Peterson, 2012) or to other measures based on it, such as Rényi entropy (e.g., Kar, Bhagat, & Routray, 2010) or Tsallis wavelet entropy (Chen & Li, 2014). Shannon's (first order) entropy only depends on the relative frequencies of the symbols in a series or a grid, without taking their organization into account. For that reason, more sophisticated measures have also been suggested, although seldom used, such as higher-order entropy or entropy rate (Porta et al., 2001). In the sequel, we use the term "entropy" in its more conventional and classical meaning of Shannon's first order information entropy.

Density can be related to (Shannon first order) entropy (Shannon & Weaver, 1949) in an exact mathematical way. In the case of a binary source such as the one used here, the entropy is a simple function of density.[2] Indeed, the entropy H of a binary source corresponding to a density d is given by

$$H = -d \times \log_2(d) - (1-d) \times \log_2(1-d)$$

This function of d is an inverse U-shaped curve with a maximum at 0.5. Friedenberg and Liby (2016) found an inverse U-shaped curve with a maximum at 0.5 when plotting participants' ratings against density. As we will see below, this non-linear relation translates in a linear positive relation between aesthetic judgments and entropy. From here on, we will use entropy instead of density, as it carries the same information as density for our purpose.

Shannon first order entropy has often been used as a measure of complexity, but as the above formula shows, it is a function of density (or symbol redundancy). As a consequence, it depends on how far from a 50% density pattern the grid is, but it does not take into account how the black cells are spread across the grid. For instance, any grid with a density of 50%, including seemingly random patterns, but also regular patterns such as a chessboard, achieves maximum entropy. Although widely used as a complexity measure, entropy has been criticized for being too narrow in scope and equivalent to a χ^2 (Gauvrit & Morsanyi, 2014).

2.2.4. Algorithmic complexity

GIF compression metrics is another natural means to assess algorithmic complexity. The algorithmic complexity of a given object is defined as the length of the shortest program that produces it and then halts (Li & Vitányi, 2008). Equivalently, it is the shortest possible compressed

[1] A possible related measure is the mean size of the parts. For a given density d in a grid of N cells, suppose that we have k parts, then, the number $d \times N/k$ is the mean size of parts, which is a possible measure of complexity linked both to the density *and* the number of parts of a given pattern.

[2] This is also true of Hartley entropy, Rényi entropy or Tsallis entropy, but not of higher-order entropy.