

The Illusion Game: A Novel Experimental Paradigm to study Visual Illusions

Dominique Makowski¹

¹ School of Social Sciences, Nanyang Technological University, Singapore

Author Note

Correspondence concerning this article should be addressed to Dominique Makowski,
HSS 04-18, 48 Nanyang Avenue, Singapore (dom.makowski@gmail.com).

Abstract

Abstract abstract.

Keywords: visual illusions, illusion game, pyllusion, illusion effect, personality

Word count: 1156

The Illusion Game: A Novel Experimental Paradigm to study Visual Illusions

Introduction

Visual illusions are fascinating stimuli capturing a key feature of our neurocognitive systems. They eloquently show that our brains did not evolve to be perfect perceptual devices, but to take into account contextual information and prior knowledge that manifests in our conscious experience (**REF**). Despite the historical interest within the fields of visual perception (**REF**), consciousness science (**REF**), or psychiatry (**REF**), several gaps remained unanswered. *mention common factor for illusions and link with disorders.*

One key challenge hindering the further development of illusion research is the relative difficulty to adapt visual illusions to an experimental setting, which typically requires the controlled modulation of the specific variables of interest. To address this issue, we first developed a parametric framework to manipulate visual illusions, that we implemented and made accessible in the open-source software *Pyllusion* (**REF**). This software allows us to generate different types of historical visual illusions with a continuous and independent modulation of two parameters, *illusion strength* and *task difficulty* (see **Figure 1**).

Indeed, many visual illusions can be seen as made of *targets* (e.g., same-length lines), which perception is biased by the *context* (e.g., in the Müller-Lyer illusion, the segments appear of different lengths when they end with inwards or outwards pointing arrows). While most of the paradigms used in the literature prompt the subjective extent the context biases the perception (“how much does the circle appear bigger”, e.g., **REFS**), *Pyllusions* allows to create illusions in which the targets are actually more or less different (i.e., one segment is longer than the other), and in which the illusion is of varying strength (the arrows are more or less pointy).

This opens the door of an experimental task in which participants have to make perceptual judgments about the targets (e.g., which segment is the longest) under different

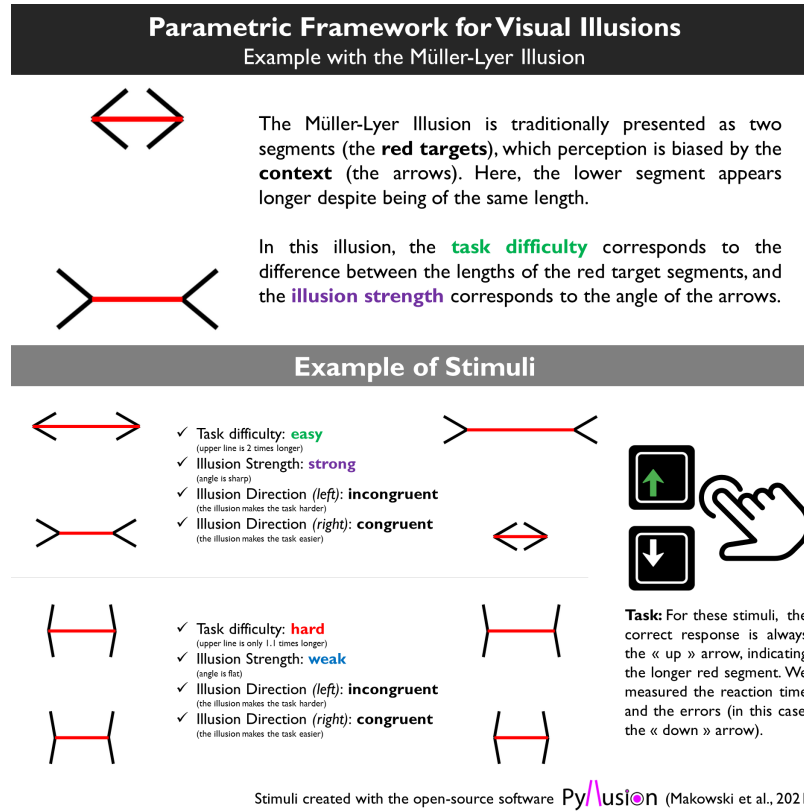


Figure 1. Explanation of the parametric framework for visual illusions (Makowski et al., 2021) applied to the Müller-Lyer illusion (above). Examples of stimuli showcasing the manipulation of the two main parameters, the task difficulty and the illusion strength (below).

conditions of objective difficulty and illusion strength. Moreover, the illusion effect can be either “incongruent” (making the task even harder by biasing the perception in the opposite way) or “congruent” (making the task easier). Although visual illusions are inherently tied to subjective perception, this framework allows a reversal of the traditional paradigm that could quantify the “objective” effect of illusions by measuring its behavioral effect (error rate and reaction times) on the performance in a perceptual task.

In the present set of preregistered studies, we will first test this novel paradigm by investigate if the effect of illusion and task difficulty can be manipulated continuously and separated statistically. Then, we will use the paradigm to assess whether 10 different historical illusions (Delboeuf, Ebbinghaus, Rod and Frame, Vertical-Horizontal, Zöllner, White, Müller-Lyer, Ponzo, Poggendorff, Contrast) share a common latent factor (a

long-standing debate). Finally, we will investigate how the the inter-individual sensitivity to illusions relates to other characteristics, such as demographic variables and personality.

In line with our full open-access standards, all the material (stimuli generation code, experiment code, raw data, analysis script with complementary figures and analyses, preregistration, etc.) is available at <https://github.com/RealityBending/IllusionGameValidation>.

Study 1

Aim

Study 1 can be seen as a pilot study which goal was to gather some preliminary data to 1) assess whether the stimuli generated by *Pyllusion* behave as expected for each of the 10 illusion types (i.e., increase of task difficulty and illusion strength leading to increased error rate) and 2) develop an intuition of the magnitude of effects, to then refine the stimuli parameters to a sensible range (i.e., not overly easy and not impossible hard).

Procedure

We generated 56 stimuli for each of the 10 illusion types. These stimuli resulted from the combination of 8 linearly-spread levels levels of task difficulty (e.g., [1, 2, 3, 4, 5, 6, 7], where 1 corresponds to the higher difficulty - i.e., the smallest objective difference between targets) and 7 levels of illusion strength (3 values of strength on the congruent side, 3 on the incongruent side, and 0; e.g., [-3, -2, -1, 0, 1, 2, 3], where the negative values correspond to congruent illusion strengths).

The 10 illusion blocks were randomly presented, and the order of the 56 stimuli within the block was also randomized. After the first series of 10 blocks, another series was done (with new randomized blocks and trials order). In total, each participant saw 56 different trials per 10 illusion type, repeated 2 times (total = 1120 trials), to which they

had to respond “as fast as possible without making errors” (i.e., an explicit double constraint to mitigate the inter-individual variability in the speed-accuracy trade off). See instructions for each illusion type in the experiment code.

Participants

Fifty-two participants were recruited via *Prolific* (REF about prolific and their data quality) with no restrictions on age or country (the only requisite was a “fluent” level of English to match the language of instructions), with a reward of about £7.5.

We removed 6 participants upon inspection of the average error rate (when close to 50%, suggesting random answers), and when the reaction time distribution was implausibly fast. For the remaining participants, we discarded blocks where the error rate was higher than 50% (possibly indicating that instructions got mixed up; e.g., participants were selecting the shorter line instead of the longer one). Finally, we removed 692 (1.37%) trials based on an implausibly short or wrong response time ($150 \text{ ms} < \text{RT} < 3000 \text{ ms}$).

The final sample included 46 participants (Mean age = 26.7, SD = 7.7, range: [19, 60]; Sex: 39.1% females, 56.5% males).

Data Analysis

The analysis of study 1 focused on the probability of errors as the main outcome variable. For each illusion, we started by visualizing the average effect of task difficulty and illusion strength to gain some intuition on the underlying generative model. Next, we tested the performance of various models differing in their specifications, such as: with or without a transformation of the task difficulty (log, square root or cubic root), with or without a 2nd order polynomial term for the illusion strength, and with or without the illusion side (up *vs.* down or left *vs.* right) as an additional predictor. We then fitted the best performing model under a Bayesian model, and compared its visualization with that

of a General Additive Model (GAM), which has an increased ability of mapping underlying potentially non-linear relationships (at the expense of model simplicity).

The analysis was carried out using *R 4.2* (**ref**), *brms* (*REF*), the *tidyverse* (*JOSS REF*) and the *easystats* collection of packages (*REFS*).

Results. The statistical models suggested that the effect of task difficulty had a cube relationship with error rate for Delboeuf and Ebbinghaus illusions (both based on circle size), square for Rod and Frame, Vertical-Horizontal, cube for Zöllner and Poggendorff, exponential for White, cube for Müller-Lyer and Ponzo (both based on lie lengths), and linear for Contrast. See details and figures in the analysis script.

Discussion

This study allowed to elaborate a more precise prior idea regarding the magnitude of the effects at stake and the type of interaction between them. It allowed us to better understand and test-out the stimuli generated by *Pyllusion*. It allowed us to uncover a few bugs and issues (for instance, the specification direction of the illusion strength was reversed in the software), which were fixed in a new release.

Crucially, this study allowed us to refine the range of task difficulty and illusion strength values in order to maximize the information gain. One notable result was the illusion effect for the Zöllner illusion, which suggested a non-linear pattern. By generating a wider range of illusion strength values, the next study will attempt at clarifying this point.

Study 2

Aim

Procedure

Perceptual decisions often exhibit a non-linear relationship with the objective physical values (e.g., a logarithmic relationship. In order to spread the difficulty levels based on the

behavioral performance (to maximize the diversity of the stimuli in regards to perceptual processes), we generated, for study 2, stimuli with the difficulty levels spaced non-linearly, depending on the best underlying model (i.e., with an exponential, square or cubic spacing if the best model was *log*, *square root* or *cube root*). For instance, a linear space of [0.1, 0.3, 0.5, 0.7] can be transformed to an exponential space of [0.1, 0.27, 0.47, 0.7].

Participants

General Discussion

Future Directions

We strongly invite researchers to explore and re-analyze our dataset with other approaches and methods to push the understanding further.

References