The Illusion Game: A Novel Experimental Paradigm Provides Evidence in Favour of a General Factor of Visual Illusion Sensitivity and Personality Correlates

Dominique Makowski[1], An Shu Te[1], Stephanie Kirk[1], Ngoi Zi Liang[1], & S.H. Annabel Chen[1, 2, 3, 4]

[1] School of Social Sciences, Nanyang Technological University, Singapore

[2] LKC Medicine, Nanyang Technological University, Singapore

[3] National Institute of Education, Singapore

[4] Centre for Research and Development in Learning, 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).

The authors made the following contributions. Dominique Makowski: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft; An Shu Te: Project administration, Resources, Investigation, Writing – original draft; Stephanie Kirk: Project administration, Resources, Writing – original draft; Ngoi Zi Liang: Project administration, Resources, Writing – original draft; S.H. Annabel Chen: Project administration, Supervision, Writing – review & editing.

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

Abstract

Visual illusions strikingly highlight how the brain uses contextual and prior information to inform our perception of reality. Unfortunately, illusion research has been hampered by the difficulty of adapting these stimuli to experimental settings, which ideally require a controlled and gradual modulation of the effects of interest. In this set of studies, we used the parametric framework for visual illusions implemented in the *Pyllusion* software to generate 10 different classic illusions (Delboeuf, Ebbinghaus, Rod and Frame, Vertical-Horizontal, Zöllner, White, Müller-Lyer, Ponzo, Poggendorff, Contrast) varying in strength. We tested the objective effect of the illusions on errors and reaction times in a perceptual discrimination task. We then extracted the participant-level performance scores (n=250), and provide evidence in favour of the existence of a general factor (labelled Factor *i*) underlying the sensitivity to different illusions. Finally, we report a positive relationship between illusion sensitivity and personality traits such as Agreeableness, Honesty-Humility, and negative relationships with Psychoticism, Antagonism, Negative Affect and Disinhibition. All the materials are available in open-access (https://github.com/RealityBending/IllusionGameValidation). We invite researchers to re-analyze the data using alternative approaches to provide complimentary findings on the effect, structure and correlates, of illusion sensitivity.

Keywords: visual illusions, illusion game, Pyllusion, personality, general factor

Word count: 1156

The Illusion Game: A Novel Experimental Paradigm Provides Evidence in Favour of a General Factor of Visual Illusion Sensitivity and Personality Correlates

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 providing veridical accounts of physical reality, but integrate prior knowledge and contextual information - blended together in our subjective conscious experience (Carbon, 2014). Despite the historical and intensive interest within the fields of visual perception (Day, 1972; Eagleman, 2001; Gomez-Villa et al., 2022), consciousness science (Caporuscio et al., 2022; Lamme, 2020), and psychiatry (Gori et al., 2016; Notredame et al., 2014; Razeghi et al., 2022; Teufel et al., 2015), several important issues remain open.

Notably, the presence of a common mechanism underlying the effect of different illusions has been contested (Cretenoud, Francis, et al., 2020; Cretenoud et al., 2019a; Hamburger, 2016; Teufel et al., 2018b); and the nature of the underlying processes - whether related to low-level features of the visual processing system (Cretenoud et al., 2019b; Gori et al., 2016) or to top-down influences of prior beliefs (Caporuscio et al., 2022; Teufel et al., 2018a) are strongly debated. The existence of dispositional or demographic correlates of illusion sensitivity - for example, higher illusion resistance has been reported in schizophrenia and autism (Giaouri & Alevriadou, 2011; Keane et al., 2014; Notredame et al., 2014; Park et al., 2022); as well as in males as compared to females (Lo & Dinov, 2011; Miller, 2001; Papageorgiou et al., 2020; Shaqiri et al., 2018) (**Is there something for personality that we could use here instead of sex?**) - is another area of controversy.

One key challenge hindering the further development of illusion research is the relative difficulty in adapting 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, which we implemented and made accessible in the open-source software *Pyllusion* (Makowski et al., 2021a). This software allows us to generate different types of classic visual illusions (e.g., Müller-Lyer, Ponzo, Delboeuf, Ebbinghaus) with a continuous and independent modulation of two parameters: *illusion strength* and *task difficulty* (see **Figure 1**).

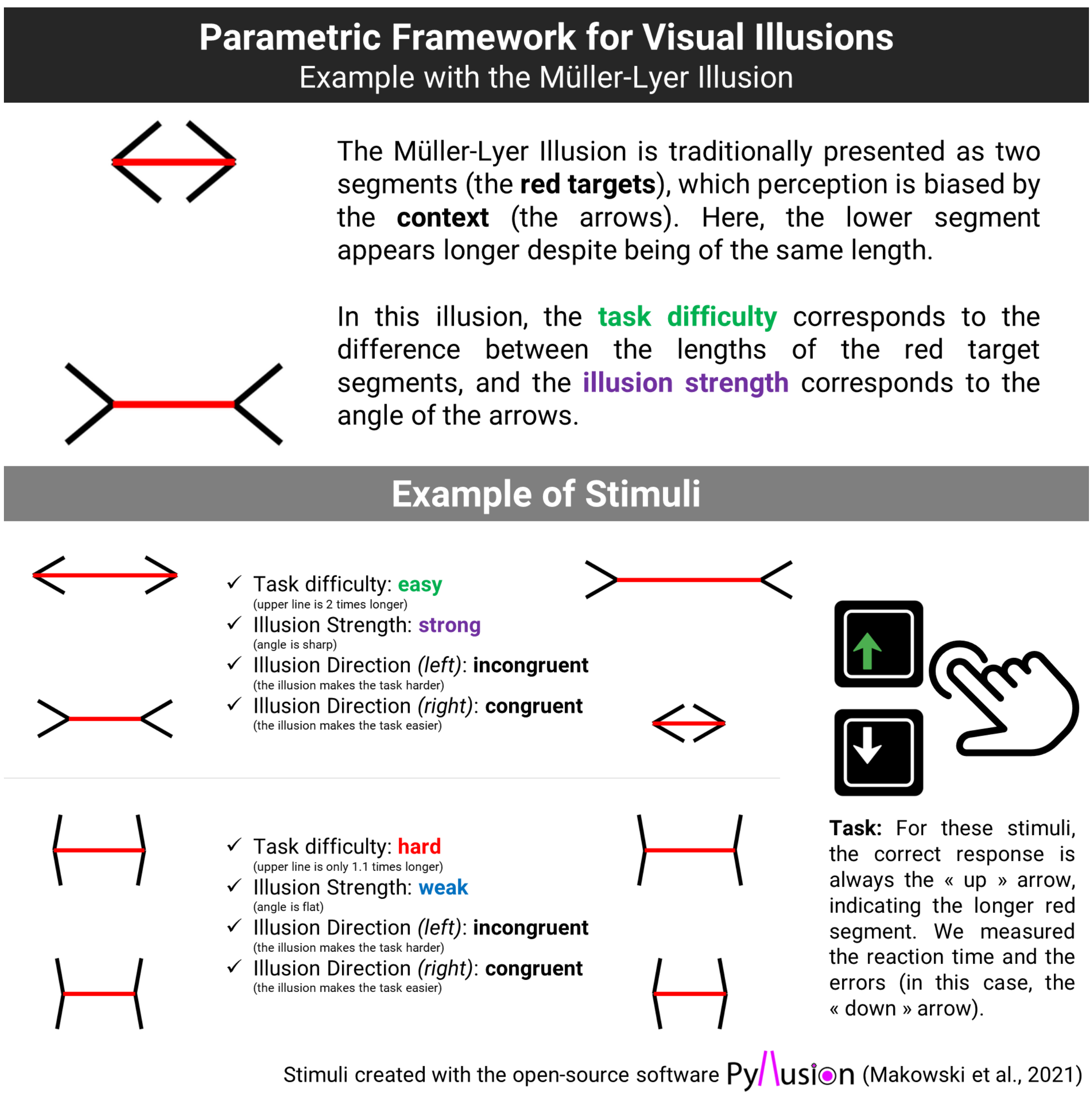

Figure 1. The parametric framework for visual illusions (Makowski et al., 2021) applied to the Müller-Lyer illusion (above). Below are examples of stimuli showcasing the manipulation of two parameters, task difficulty and illusion strength.

Indeed, many visual illusions can be seen as being composed of *targets* (e.g., same-length lines), of which perception is biased by the *context* (e.g., in the Müller-Lyer illusion, the same-length line segments appear to have different lengths when they end with inwards or outwards pointing arrows). Past illusion studies traditionally employed paradigms focusing on participants' subjective experience, by asking them to what extent they perceive two identical targets as different (Lányi et al., 2022), or having them adjust the targets to a reference stimulus relying only on their perception (Grzeczkowski et al., 2018; Mylniec & Bednarek, 2016a). Alternatively, *Pyllusion* allows the creation of illusions in which the targets are objectively different (e.g., one segment is truly more or less longer than the other), and in which the illusion varies in strength (the biasing angle of the arrows is more or less acute).

This opens the door for an experimental task in which participants make perceptual judgments about the targets (e.g., which segment in the longest) under different conditions of objective difficulty and illusion strength. Moreover, the illusion effect can be either "incongruent" (making the task more difficult 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 to potentially 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 investigating if the effect of illusion and task difficulty can be manipulated continuously, and separately modeled statistically. Then, we will further utilize the paradigm to assess whether 10 different classic illusions (Delboeuf, Ebbinghaus, Rod and Frame, Vertical-Horizontal, Zöllner, White, Müller-Lyer, Ponzo, Poggendorff, Contrast) share a common latent factor. Finally, we will investigate how the the inter-individual sensitivity to illusions relates to dispositional variables, such as demographic characteristics and personality.

In line with open-science 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 experiment aiming to gather some preliminary data to assess if the stimuli generated by *Pyllusion* behaves as expected for each of the 10 illusion types (i.e., whether an increase of task difficulty and illusion strength leads to an increase of errors); and develop an intuition about the magnitude of effects, to refine the stimuli parameters to a more sensible range (i.e., not overly easy and not impossibly hard) for the next study.

Procedure

We generated 56 stimuli for each of the 10 illusion types. These stimuli resulted from the combination of 8 linearly-spread levels of task difficulty (e.g., [1, 2, 3, 4, 5, 6, 7], where 1 corresponds to the highest 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 negative values correspond to congruent illusion strengths).

The 10 illusion blocks were randomly presented, and the order of the 56 stimuli within the blocks was also randomized. After the first series of 10 blocks, another series was done (with new randomized order of blocks and trials). 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). The task was implemented using *jsPsych* (De Leeuw, 2015). The instructions for each illusion type are available in the experiment code.

Participants

Fifty-two participants were recruited via *Prolific* (www.prolificacademic.co.uk), a crowd-sourcing platform providing high data quality (Peer et al., 2022). The only inclusion criterion was a fluent proficiency in English to ensure that the task instructions would be well-understood. Participants were incentivised with a reward of about £7.5 for completing the task, which took about 50 minutes to finish.

We removed 6 participants upon inspection of the average error rage (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 misunderstood; 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 long response time (< 150 ms or > 3000 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 logistic 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 framework, and compared its visualization with that of a General Additive Model (GAM), which has an increased ability of mapping underlying potential non-linear relationships (at the expense of model simplicity).

The analysis was carried out using *R 4.2* (R Core Team, 2022), *brms* (Bürkner, 2017), the *tidyverse* (Wickham et al., 2019), and the *easystats* collection of packages (Lüdecke et al., 2021, 2019; Makowski et al., 2020; Makowski, Ben-Shachar, & Lüdecke, 2019).

Results

The statistical models suggested that the effect of task difficulty had a cubic relationship with error rate for the Delboeuf and Ebbinghaus illusions (both composed of circular shapes), square relationship for the Rod and Frame and Vertical-Horizontal illusions, cubic relationship for the Zöllner and Poggendorff illusions, exponential relationship for the White illusion, cubic relationship for the Müller-Lyer and Ponzo illusions (both based on line lengths), and linear relationship for the Contrast illusion. All models suggested a significant effect of illusion strength and task difficulty. See details and figures in the analysis script.

Discussion

This study provided a clearer understanding of the magnitude of the parametric effects at stake and the type of interaction between them. Furthermore, it allowed us to better understand and test the stimuli generated by *Pyllusion*, as well as uncover technical bugs and issues (for instance, the specification direction of the illusion strength was reversed for a few illusions), which were fixed by a new software release. Crucially, this study allowed us to refine the range of task difficulty and illusion strength values in order to maximize information gain.

In most illusions, the task difficulty exhibited monotonic power-law scaled effects, which is in line with the psychophysics literature on perceptual decisions (Bogacz et al., 2006; Ditzinger, 2010; Shekhar & Rahnev, 2021). One notable result was the illusion effect pattern for the Zöllner illusion, which suggested a non-linear relationship. By generating a wider range of illusion strength values, the next study will attempt at clarifying this point.

Study 2

Aim

The aim of study 2 was two-fold. In the first part, we carefully modeled the error rate and the reaction time of each illusion type in order to validate our novel paradigm and show that the effect of illusions can be manipulated continuously. In the second part, we derived the participant-level scores from the models (i.e., the effect of illusion strength for each individual) and analyzed their latent factors structure.

Procedure

The paradigm of study 2 was similar to that of study 1, with the following changes. The illusory stimuli were re-generated within a refined space of parameters based on the results of study 1. Moreover, taking into account the findings of study 1, we used non-linearly spaced difficulty levels, depending on the best underlying model (i.e., with an exponential, square or cubic spacing depending on the relationship). For instance, a linear space of [0.1, 0.4, 0.7, 1.0] can be transformed to an exponential space of [0.1, 0.34, 0.64, 1.0].

Additionally, instead of repeating each stimulus two times, we generated illusions using more levels of difficulty and illusion strength. As such, for each illusion type, we generated a total of 134 stimuli that were split into two groups (67 stimuli per illusion block). Furthermore, instead of a simple break screen, we added two personality questionnaires between the two series of 10 illusion blocks (see study 3).

Participants

Using the same recruitment procedure as in study 1, we recruited 256 participants, out of which 6 were identified as outliers and excluded, leaving a final sample of 250 participants (Mean age = 26.5, SD = 7.6, range: [18, 69]; Sex: 48% females, 52% males). Please see study 3 for the full demographic breakdown. We discarded blocks with more than 50% of errors (2.16% of trials) and 0.76% trials with extreme response times (< 125 ms or > 4 SD above mean).

Data Analysis

The first part of the analysis focused on modelling the effect of illusion strength and task difficulty on errors and reaction time (RT), within each illusion. In order to achieve that, we started by fitting General Additive Models (GAMs), which can accommodate possible non-linear effects and interactions. Errors were analyzed using Bayesian logistic mixed models, and RTs of correct responses were analyzed using an ex-Gaussian family with the same fixed effects entered for the location μ (mean), scale σ (spread) and tail-dominance τ of the RT distribution (Balota & Yap, 2011; Matzke & Wagenmakers, 2009).

Using GAMs as the "ground-truth" models, we attempted at approximating them using general linear models, which have the advantage of estimating the participant-level variability of the effects (via random slopes). Following a comparison of models with a combination of transformations (raw, log, square root or cubic root) on the main predictors (task *difficulty* and illusion *strength*), we selected and fitted the best model (best on their indices of fit), and compared their output visually (see **Figure 2**).

We then extracted the inter-individual variability in the effect of illusion strength and its interaction with task difficulty, and used it as participant-level scores. Finally, We explored the relationship of these indices across different illusions using exploratory factor analysis (EFA) and structural equation modelling (SEM).

Results

The best models were $log(diff) * strength$ for Delboeuf; $sqrt(diff) * strength$ for Ebbinghaus; $log(diff) * log(strength)$ for Rod and Frame; $sqrt(diff) * sqrt(strength)$ for Vertical-Horizontal; $cbrt(diff) * strength$ for Zöllner; $diff * sqrt(strength)$ and $log(diff) * strength$ respectively for errors and RT in White; $sqrt(diff) * sqrt(strength)$ and $sqrt(diff) * strength$ respectively for errors and RT in Müller-Lyer; $cbrt(diff) * strength$ for Ponzo; $cbrt(diff) * sqrt(strength)$ and $cbrt(diff) * strength$ respectively for errors and RT in Poggendorff; $sqrt(diff) * sqrt(strength)$ for Contrast. In all of these models, the effects of illusion strength, task difficulty and their interaction were significant.

For errors, most of the models closely matched their GAMs counterpart (see **Figure 2**), with the exception of Delboeuf (for which the GAM suggested a non-monotonic effect of illusion strength with a local minimum at 0) and Zöllner (for which theoretically congruent illusion effects were related to increased error rate).

For RTs, the GAMs suggested a consistent non-linear relationship between RT and illusion strength: as the illusion strength increase beyond a certain threshold, the participants respond faster. While this is is not surprising (strong illusions are likely so effective in biasing perception that it is "easier", i.e., faster, to make the wrong decision), the linear models were not designed to capture this - likely quadratic - pattern and hence are not good representatives of the underlying dynamics. As such, we decided not to use them for the individual scores analysis.

Though imperfect, we believe that the random-slope models capture inter-individual differences with more accuracy (and are also more conservative estimates due shrinkage) than basic empirical scores, such as the total number of errors, or the average RT. Thus, for

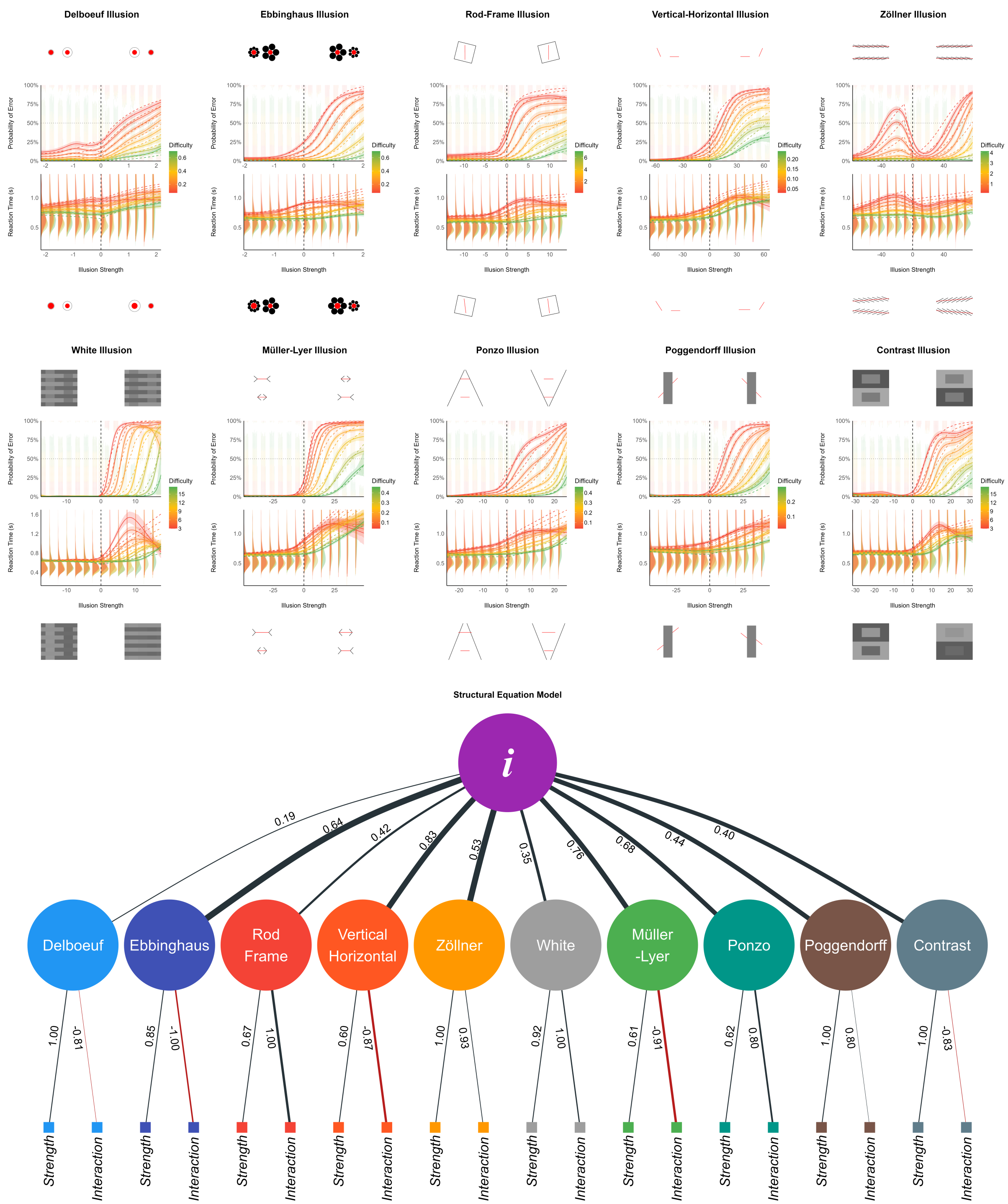

Figure 2. Top: the effect of illusion strength and task difficulty on the error rage and reaction time (RT) for each individual illusions. The solid line represent the General Additive Model (GAM), and the dashed line correspond to its approximation via linear models. Descriptive data is shown with stacked dots (errors are hanging from the top) and distributions for RTs. Negative values for illusion strength correspond to congruent (i.e., facilitating) illusion effects. Task difficulty (the objective difference between the targets of perceptual decision) levels are shown as colors, with lowest values corresponding to harder trials. Each illusion type is surrounded by 4 extreme examples of stimuli, corresponding to the hardest difficulty (on top) and the strongest illusion (on the right for incongruent illusions). Bottom: We extracted the effect slope of the illusion strength and its interaction with task difficulty for each participant. We fitted a Structural Equation Model (SEM) suggesting that these manifest variables group to first-level illusion-specific latent factors, which then load on a general factor of illusion sensitivity (Factor i).

each illusion and within each participant, we extracted the effect of illusion strength and its interaction with task difficulty when the illusion effect was incongruent. These twenty participant-level scores were subjected to exploratory factor analysis (EFA). The Method Agreement Procedure (Lüdecke et al., 2020) suggested the presence of 7 latent factors. An oblique (*oblimin* rotation) factor solution explaining 66.69% of variance suggested separate dimensions for the effect of Zöllner, White, Poggendorff, Contrast, Ebbinghaus, Delboeuf, and a common factor for the parameters related to Müller-Lyer, Vertical-Horizontal, Ponzo and Rod and Frame. We submitted these factors to a second-level analysis and extracted two orthogonal (*varimax* rotation) factors. The first factor was loaded by all the previous dimensions with the exception of Delboeuf, which formed its own separate factor.

Finally, we tested this data-driven model (*m0*) against four other structural models using structural equation modelling (SEM): one in which the two parameters of each of the 10 illusions (illusion strength and interaction with task difficulty) loaded on separate factors, which then all loaded on a common factor (*m1*); one which the parameters were grouped by illusion type (lines, circles, contrast and angle) before loading on a common factor (*m2*); one in which all the parameters related to strength, and all the parameters related to the interaction loaded onto two respective factors, which then loaded on a common factor (*m3*); and one in which there was no intermediate level: all 20 parameters loaded directly on a common factor (*m4*).

The model *m1*, in which the parameters loaded on a first level of 10 illusion-specific factors, which then all loaded on a common factor significantly outperformed the other models. Its indices of fit were ranging from acceptable to satisfactory (CFI = .92; SRMR = .08; NNFI = .91; PNFI = .74; RMSEA = .08), and all the specified effects were significant. The illusion-specific latent factors were loaded positively by the sensitivity to illusion strength, and positively by the interaction effect with task difficulty (with the exception of Delboeuf, Ebbinghaus, Vertical-Horizontal, Müller-Lyer and Contrast, for which the loading was negative). The general factor of illusion sensitivity, labelled Factor i (i- for illusion), explained 48.02% of the total variance of the initial dataset, and was strongly related to Vertical-Horizontal ($\beta_{std.} = 0.83$), Müller-Lyer ($\beta_{std.} = 0.76$), Ponzo ($\beta_{std.} = 0.65$), Ebbinghaus ($\beta_{std.} = 0.64$); moderately to Zöllner ($\beta_{std.} = 0.53$), Poggendorff ($\beta_{std.} = 0.44$), Rod and Frame ($\beta_{std.} = 0.42$), Contrast ($\beta_{std.} = 0.40$) and White ($\beta_{std.} = 0.35$); and weakly to Delboeuf ($\beta_{std.} = 0.19$). We then computed, for each participant, its score for the 10 illusion-specific factors and for the general Factor i.

We have to keep in mind that these individual scores are the result of several layers of simplification: 1) the individual coefficient is that of simpler models that sometimes do not perfectly capture the underlying dynamics (especially in the case of Delboeuf and Zöllner); 2) we only used the models on error rate, which could be biased by the speed-accuracy decision criterion used by participants; 3) the structural equation model used to compute the scores also incorporated multiple levels of abstractions. Thus, in order to validate the individual scores, we computed the correlation between them and simple empirical scores, such as the average error rate and the mean RT in the task. This analysis revealed strong and significant correlations between each illusion-specific factor and the average amount of errors in its respective task. Moreover, each individual score was strongly associated with the average RT across multiple illusion types. This suggests that the individual scores obtained from the structural equation model do capture the sensitivity of each participant to visual illusions, manifesting in both the number of errors and high reaction times.

Discussion

This study confirmed that it was possible to continuously manipulate the effect of illusion strength for 10 classical illusions. Increasing the illusion strength increased the likelihood of errors, as well as the average and spread of RTs (but only up to a point, after which participants become faster at responding with the wrong answer). Future studies are needed to explore reaction times and try to identify the most appropriate models, and / or use models that integrate errors and reaction time (e.g., drift diffusion models).

The effect on errors was monotonic for most illusions, with the exception of Delboeuf and Zöllner. For both of them, mildly congruent illusion strengths (which theoretically were supposed to be associated with less errors than incongruent effects) were related to small and strong increases of errors, respectively. For the Delboeuf illusion, we believe that this was due to an artifact caused by the illusion generation algorithm: the outline of the target circles was always created as slightly bigger, which made the difference between them more obvious at an illusion strength of 0. This was fixed in latest release of *Pyllusion* (v1.2), which now generate outlines of the same size as the target circle. For the Zöllner illusion, the observed non-monotonic pattern is actually consistent with previous reports (Kitaoka, 2007; Kitaoka & Ishihara, 2000), suggesting an acute angle contraction effect at very small as well as at sufficiently large angles (below 10 degrees for the former and between 50 to 90 degrees for the latter) between the target horizontal line and the biasing horizontal bars when the illusion strength is weak.

Finally, this study provided evidence for both the existence of illusion-specific factors, as well as for a common latent factor (labelled Factor *i*) that explained about half of the total variance. These participant-level scores were positively related to the error rate and average reaction time, and can thus be interpreted as indices of illusion sensitivity.

Study 3

Aim

Study 3 aimed at investigating the links between the inter-individual scores of illusion sensitivity (obtained in study 2), and demographic and dispositional variables. **TODO: Insert a bit of literature about why and some findings justifying what we did**.

Procedure

This study was based on the data collected in study 2. The variables of interest here were taken from the questionnaires that were inserted in between the two series of illusion blocks. We used the *IPIP6* (24 items, Sibley et al., 2011) to measure 6 "normal" personality traits (Extraversion, Openness, Conscientiousness, Agreeableness, Neuroticism and Honesty-humility), and the PID-5 (25 items, Hopwood et al., 2012) to measure "pathological" personality traits (Disinhibition, Antagonism, Detachment, Negative Affect and Psychoticism). The participants were the same as in study 2 (see **Figure 3**). However, due to a technical issue, no personality data was recorded for the first eight participants.

Data Analysis

For each of the individual illusion sensitivity scores (10 illusion-specific factors and the general Factor i), we tested the effect of contextual variables (screen size, screen refresh rate), demographic variables (sex, education, age) and personality. As the supplementary material contains the detailed results, we will here only report the significant results (based on the Bayes Factor BF or the Probability of Direction *pd*, see Makowski, Ben-Shachar, Chen, et al., 2019).

Results

The Bayesian correlation analysis (with narrow priors centered around a null effect) between the illusion scores and contextual variables (screen size and refresh rate) provided weak evidence in favor of an absence of effect, with the exception of the two contrast-based illusions. Anecdotal ($BF_{10} = 2.05$) and moderate evidence ($BF_{10} = 4.11$) was found for a negative correlation between screen size and the sensitivity to the White and the Contrast illusion, respectively. To test whether this result could be an artifact related to the highly skewed screen size distribution (caused by very few participants with extreme screen sizes), we re-ran a robust correlation (with rank-transformed values), which provided even

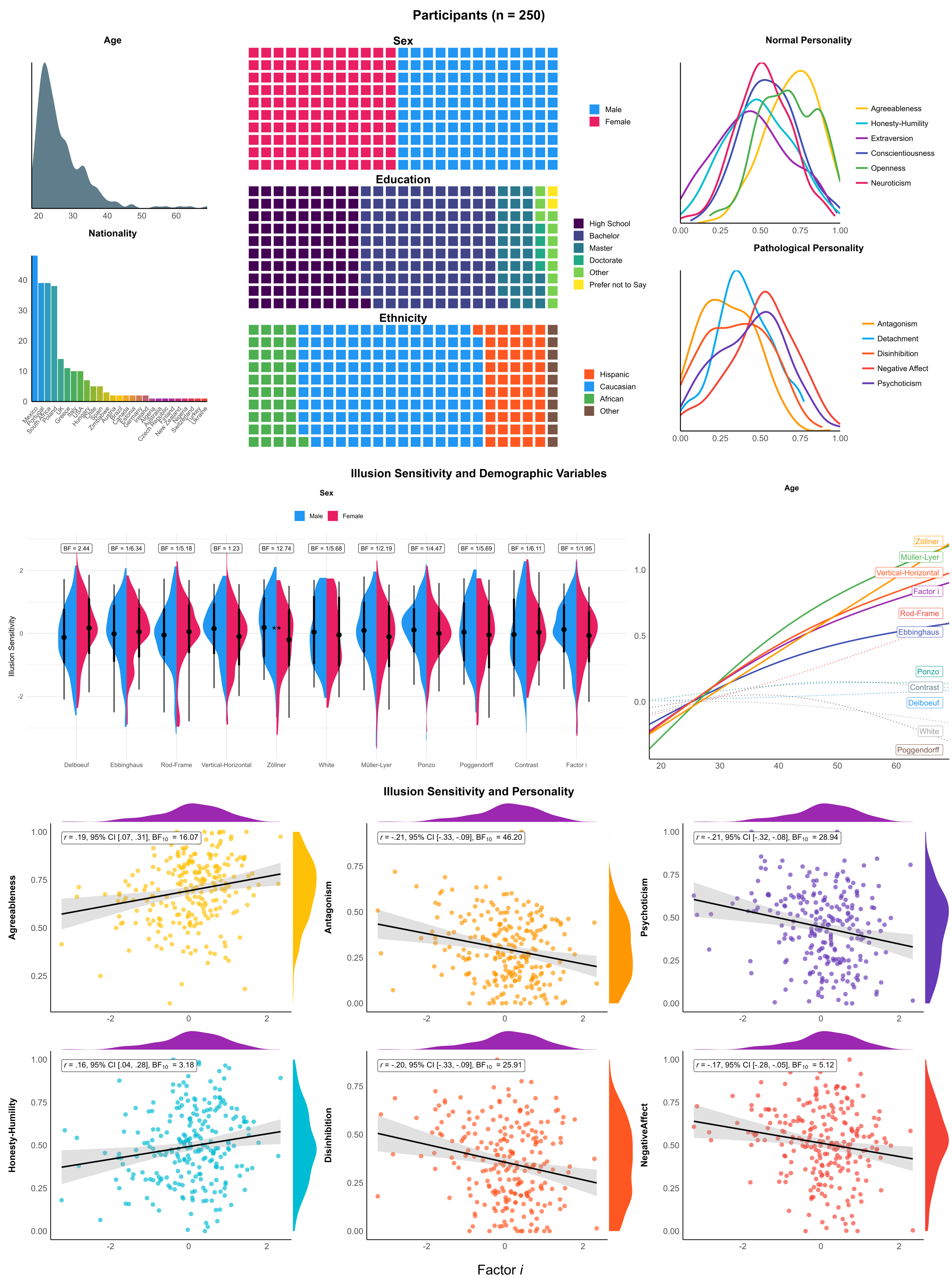

Figure 3. The upper plots show the distribution of demographic and dispositional variables. The middle plots show the relationship between illusion sensitivty scores, sex, and age (solid lines indicate significant relationships). Bottom plots show the correlation between the general factor of illusion sensitivity (Factor i), and personality traits.

stronger evidence in favor of the effect existence ($BF_{10} = 28.19$, $BF_{10} = 4.31$ for White and Contrast, respectively).

The Bayesian t-tests on the effect of sex suggested anecdotal to moderate evidence in favour of the null effect for all scores, with the exception of the sensitivity to the Zöllner illusion, which was higher in males as compared to females ($\Delta = -0.37$, 95% CI [-0.62, -0.13], $BF_{10} = 12.74$). We fitted Bayesian linear models with the education level entered as a monotonic predictor (appropriate for ordinal variables, Bürkner & Charpentier, 2020), which yielded no significant effects. For age, we fitted two types of models for each score, one general additive models (GAM) and a 2nd order polynomial model. These consistently suggested a significant positive linear relationship between age and Factor i ($pd = 100\%$), as well as the sensitivity to Müller-Lyer ($pd = 100\%$), Vertical-Horizontal ($pd = 100\%$), Zöllner ($pd = 100\%$) and Ebbinghaus ($pd = 99\%$) illusions.

Regarding "normal" personality traits, Bayesian correlations suggested substantial evidence in favor of a positive relationship between *Honesty-Humility* and Zöllner ($BF_{10} > 100$), Vertical-Horizontal ($BF_{10} = 9.78$) and the Factor i ($BF_{10} = 4.00$); as well as between *Agreeableness* and Vertical-Horizontal ($BF_{10} = 25.06$), Ponzo ($BF_{10} = 4.88$) and the Factor i ($BF_{10} = 19.65$).

Regarding "pathological" personality traits, the results yielded strong evidence in favor of a negative relationship between multiple illusion scores and multiple traits. *Antagonism* was associated with the sensitivity to Vertical-Horizontal ($BF_{10} > 100$), Müller-Lyer ($BF_{10} = 21.57$), Ponzo ($BF_{10} = 17.97$) illusions, and the Factor i ($BF_{10} = 55.45$); *Psychoticism* was associated with the sensitivity to Vertical-Horizontal ($BF_{10} = 66.63$) and Müller-Lyer ($BF_{10} = 35.59$) illusions, and the Factor i ($BF_{10} = 35.02$); *Disinhibition* was associated with the sensitivity to Vertical-Horizontal ($BF_{10} = 25.38$), Zöllner ($BF_{10} = 7.59$), Müller-Lyer ($BF_{10} = 5.89$) illusions, and the Factor i ($BF_{10} = 31.42$); and *Negative Affect* was associated with Zöllner ($BF_{10} = 62.04$), Vertical-Horizontal ($BF_{10} = 12.65$), Müller-Lyer ($BF_{10} = 3.17$), and the Factor i ($BF_{10} = 6.39$). The last remaining trait, *Detachment*, did not share any relationship with illusion sensitivity.

Discussion

Despite the widespread interest in finding associations between personality correlates and a general sensitivity to visual illusions, the supporting literature remains relatively mixed. Whereas some researchers report no relationship between inter-individual traits and individuals' susceptibility to illusions (Cretenoud, Grzeczkowski, et al., 2020; Grzeczkowski et al., 2017), others have found pathological traits such as higher Aggression-Hostility and Narcissism, and lower Impulsive Sensation-Seeking to be significantly correlated to a greater resistance against such illusory effects (Ohmann & Burgmer, 2016; Zhang et al., 2017). Furthermore, research on patients with schizophrenia have also generally demonstrated lower illusion sensitivity to be linked to symptoms of the disorder, such as greater hostility as well as social and emotional withdrawal (Pessoa et al., 2008).

General Discussion

Using the parametric illusion generation framework we developed, *Pyllusion* (Makowski et al., 2021b), we have hence shown that illusions can be manipulated continuously across several different visual illusions. This opens the door for new illusions-based paradigms and tasks, therefore making it possible for future researchers to directly manipulate specific features and parameters of the illusion that are of interest. The validation of this novel framework also affords future illusion scientists a standardized measure of illusion susceptibility, instead of relying on conventional methods that depend upon participants' subjective perceptions. In our paradigm, in which we apply this approach to a reaction-time task, we were able to measure inter-individual scores of objective illusion sensitivity.

Notably, the general sensitivity to illusions Factor i was negatively associated with *Antagonism*, *Psychoticism*, *Disinhibition* and *Negative Affect*.

Most notably, there is currently no universally agreed upon neurocognitive mechanism that explains individuals' susceptibility to visual illusions (Mylniec & Bednarek, 2016b). For instance, while some researchers have tried to explain our sensitivity to illusory effects as a result of deficits in the low-level visual processing system (Cretenoud et al., 2019b; Gori et al., 2016), others have provided a compelling case using a top-down approach, suggesting that such visual phenomena occur as a result of a conflict between our visual input and our prior beliefs(Caporuscio et al., 2022; Teufel et al., 2018a).

Furthermore, results from studies that have been conducted to elucidate the mechanism underlying our susceptibility towards visual illusions remain relatively mixed. Whereas higher resistance towards such illusions have been reported for individuals with pathologically strong prior beliefs (such as schizophrenics) and atypical sensory perception (for example, those with autism spectrum disorder [ASD]) (Giaouri & Alevriadou, 2011; Keane et al., 2014; Notredame et al., 2014; Park et al., 2022), other studies have found no significant differences between such individuals and healthy controls (Kaliuzhna et al., 2019; Spencer & Ghorashi, 2014; Tibber et al., 2013; Yang et al., 2012) or only a weak correlation between the magnitude of visual illusions and such individuals' susceptibility to these illusory effects (Grzeczkowski et al., 2018; Manning et al., 2017).

Future Directions

We strongly invite researchers to explore and re-analyze our dataset with other approaches and methods to push the understanding of visual illusions and illusion sensitivity further. The task, data and analysis script are available in open-access at **https://github.com/RealityBending/IllusionGameValidation**.

Acknowledgments

We would like to thank Tam Pham and Zen J. Lau for their contribution to *Pyllusion*, as well as Prof Dólos for the inspiration.

References

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science*, *20*(3), 160–166.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C., & Charpentier, E. (2020). Modelling monotonic effects of ordinal predictors in bayesian regression models. *British Journal of Mathematical and Statistical Psychology*, *73*(3), 420–451.

Caporuscio, C., Fink, S. B., Sterzer, P., & Martin, J. M. (2022). When seeing is not believing: A mechanistic basis for predictive divergence. *Consciousness and Cognition*, *102*, 103334. https://doi.org/10.1016/j.concog.2022.103334

Carbon, C.-C. (2014). Understanding human perception by human-made illusions. *Frontiers in Human Neuroscience*, *8*. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00566

Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision*, *20*(8), 12. https://doi.org/10.1167/jov.20.8.12

Cretenoud, A. F., Grzeczkowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, *20*(6), 4. https://doi.org/10.1167/jov.20.6.4

Cretenoud, A. F., Karimpur, H., Grzeczkowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019b). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, *19*(14), 12. https://doi.org/10.1167/19.14.12

Cretenoud, A. F., Karimpur, H., Grzeczkowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019a). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, *19*(14), 12. https://doi.org/10.1167/19.14.12

Day, R. H. (1972). Visual Spatial Illusions: A General Explanation: A wide range of visual illusions, including geometrical distortions, can be explained by a single principle. *Science*, *175*(4028), 1335–1340. https://doi.org/10.1126/science.175.4028.1335

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12.

Ditzinger, T. (2010). *Optical illusions: Examples for nonlinear dynamics in perception* (Vol. 328, pp. 179–191).

Eagleman, D. M. (2001). Visual illusions and neurobiology. *Nature Reviews Neuroscience*, *2*(12), 920–926. https://doi.org/10.1038/35104092

Giaouri, S., & Alevriadou, A. (2011). Are children with down syndrome susceptible to visual illusions? *Procedia - Social and Behavioral Sciences*, *15*, 1988–1992. https://doi.org/10.1016/j.sbspro.2011.04.040

Gomez-Villa, A., Martín, A., Vazquez-Corral, J., Bertalmío, M., & Malo, J. (2022). On the synthesis of visual illusions using deep generative models. *Journal of Vision*, *22*(8), 2. https://doi.org/10.1167/jov.22.8.2

Gori, S., Molteni, M., & Facoetti, A. (2016). Visual illusions: An interesting tool to investigate developmental dyslexia and autism spectrum disorder. *Frontiers in Human Neuroscience*, *10*, 175. https://doi.org/10.3389/fnhum.2016.00175

Grzeczkowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292. https://doi.org/10.1016/j.visres.2016.10.006

Grzeczkowski, L., Roinishvili, M., Chkonia, E., Brand, A., Mast, F. W., Herzog, M. H., & Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939. https://doi.org/10.1016/j.psychres.2018.10.063

Hamburger, K. (2016). Visual Illusions Based on Processes: New Classification System Needed: *Perception*. https://doi.org/10.1177/0301006616629038

Hopwood, C. J., Thomas, K. M., Markon, K. E., Wright, A. G. C., & Krueger, R. F. (2012). DSM-5 personality traits and DSM–IV personality disorders. *Journal of Abnormal Psychology*, *121*(2), 424–432. https://doi.org/10.1037/a0026656

Kaliuzhna, M., Stein, T., Rusch, T., Sekutowicz, M., Sterzer, P., & Seymour, K. J. (2019). No evidence for abnormal priors in early vision in schizophrenia. *Schizophrenia Research*, *210*, 245–254. https://doi.org/10.1016/j.schres.2018.12.027

Keane, B. P., Joseph, J., & Silverstein, S. M. (2014). Late, not early, stages of Kanizsa shape perception are compromised in schizophrenia. *Neuropsychologia*, *56*, 302–311. https://doi.org/10.1016/j.neuropsychologia.2014.02.001

Kitaoka, A. (2007). Tilt illusions after Oyama (1960): A review [1]: Tilt illusions: A review. *Japanese Psychological Research*, *49*(1), 7–19. https://doi.org/10.1111/j.1468-5884.2007.00328.x

Kitaoka, A., & Ishihara, M. (2000). Three elemental illusions determine the zöllner illusion. *Perception & Psychophysics*, *62*, 569–575.

Lamme, V. A. F. (2020). Visual functions generating conscious seeing. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00083

Lányi, O., Keri, S., Pálffy, Z., & Polner, B. (2022). *Can you believe your eyes? Positive schizotypy is associated with increased susceptibility to the müller-lyer illusion.*

Lo, K. C., & Dinov, D. I. (2011). *Investigation of Optical Illusions on the Aspects of Gender and Age*. *24*, 8.

Lüdecke, D., Ben-Shachar, M., Patil, I., & Makowski, D. (2020). Extracting, computing and exploring the parameters of statistical models using R. *Journal of Open Source Software*, *5*(53), 2445. https://doi.org/10.21105/joss.02445

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Lüdecke, D., Waggoner, P., & Makowski, D. (2019). Insight: A unified interface to access information from model objects in R. *Journal of Open Source Software*, *4*(38), 1412. https://doi.org/10.21105/joss.01412

Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdecke, D. (2019). Indices of effect existence and significance in the bayesian framework. *Frontiers in Psychology*, *10*, 2767.

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Makowski, D., Ben-Shachar, M., Patil, I., & Lüdecke, D. (2020). Methods and algorithms for correlation analysis in R. *Journal of Open Source Software*, *5*(51), 2306. https://doi.org/10.21105/joss.02306

Makowski, D., Lau, Z. J., Pham, T., Paul Boyce, W., & Annabel Chen, S. H. (2021a). A Parametric Framework to Generate Visual Illusions Using Python. *Perception*, *50*(11), 950–965. https://doi.org/10.1177/03010066211057347

Makowski, D., Lau, Z. J., Pham, T., Paul Boyce, W., & Annabel Chen, S. H. (2021b). A Parametric Framework to Generate Visual Illusions Using Python. *Perception*, *50*(11), 950–965. https://doi.org/10.1177/03010066211057347

Manning, C., Morgan, M. J., Allen, C. T. W., & Pellicano, E. (2017). Susceptibility to ebbinghaus and müller-lyer illusions in autistic children: A comparison of three different methods. *Molecular Autism*, *8*(1), 16. https://doi.org/10.1186/s13229-017-0127-y

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-gaussian and shifted wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, *16*(5), 798–817.

Miller, R. J. (2001). Gender differences in illusion response: The influence of spatial strategy and sex ratio. *Sex Roles*, *44*(3/4), 209–225.

https://doi.org/10.1023/A:1010907204430

Mylniec, A., & Bednarek, H. (2016b). Field dependence, efficiency of information processing in working memory and susceptibility to orientation illusions among architects. *Polish Psychological Bulletin*, *47*(1), 112–122. https://doi.org/10.1515/ppb-2016-0012

Mylniec, A., & Bednarek, H. (2016a). Field dependence, efficiency of information processing in working memory and susceptibility to orientation illusions among architects. *Polish Psychological Bulletin*, *47*(1), 112–122. https://doi.org/10.1515/ppb-2016-0012

Notredame, C.-E., Pins, D., Deneve, S., & Jardri, R. (2014). What visual illusions teach us about schizophrenia. *Frontiers in Integrative Neuroscience*, *8*, 63. https://doi.org/10.3389/fnint.2014.00063

Ohmann, K., & Burgmer, P. (2016). Nothing compares to me: How narcissism shapes comparative thinking. *Personality and Individual Differences*, *98*, 162–170. https://doi.org/10.1016/j.paid.2016.03.069

Papageorgiou, C., Stachtea, X., Papageorgiou, P., Alexandridis, A. T., Makris, G., Chrousos, G., & Kosteletos, G. (2020). Gender-dependent variations in optical illusions: evidence from N400 waveforms. *Physiological Measurement*, *41*(9), 095006. https://doi.org/10.1088/1361-6579/abb2eb

Park, S., Zikopoulos, B., & Yazdanbakhsh, A. (2022). Visual illusion susceptibility in autism: A neural model. *European Journal of Neuroscience*, *56*. https://doi.org/10.1111/ejn.15739

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Pessoa, V. F., Monge-Fuentes, V., Simon, C. Y., Suganuma, E., & Tavares, M. C. H. (2008). The miiller-lyer illusion as a tool for schizophrenia screening. *Reviews in the*

Neurosciences, *19*(2-3). https://doi.org/10.1515/REVNEURO.2008.19.2-3.91

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Razeghi, R., Arsham, S., Movahedi, A., & Sammaknejad, N. (2022). The effect of visual illusion on performance and quiet eye in autistic children. *Early Child Development and Care*, *192*(5), 807–815. https://doi.org/10.1080/03004430.2020.1802260

Shaqiri, A., Roinishvili, M., Grzeczkowski, L., Chkonia, E., Pilz, K., Mohr, C., Brand, A., Kunchulia, M., & Herzog, M. H. (2018). Sex-related differences in vision are heterogeneous. *Scientific Reports*, *8*(1), 7521. https://doi.org/10.1038/s41598-018-25298-8

Shekhar, M., & Rahnev, D. (2021). The nature of metacognitive inefficiency in perceptual decision making. *Psychological Review*, *128*(1), 45–70. https://doi.org/10.1037/rev0000249

Sibley, C., Luyten, N., Wolfman, M., Mobberley, A., Wootton, L. W., Hammond, M., Sengupta, N., Perry, R., West-Newman, T., Wilson, M., McLellan, L., Hoverd, W. J., & Robertson, A. (2011). The mini-IPIP6: Validation and extension of a short measure of the big-six factors of personality in new zealand. *New Zealand Journal of Psychology*, *40*, 142–159.

Spencer, K., & Ghorashi, S. (2014). Oscillatory dynamics of gestalt perception in schizophrenia revisited. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00068

Teufel, C., Dakin, S. C., & Fletcher, P. C. (2018a). Prior object-knowledge sharpens properties of early visual feature-detectors. *Scientific Reports*, *8*(1), 10853. https://doi.org/10.1038/s41598-018-28845-5

Teufel, C., Dakin, S., & Fletcher, P. (2018b). Prior object-knowledge sharpens properties of early visual feature-detectors. *Scientific Reports*, *8*. https://doi.org/10.1038/s41598-018-28845-5

Teufel, C., Subramaniam, N., Dobler, V., Perez, J., Finnemann, J., Mehta, P. R., Goodyer, I. M., & Fletcher, P. C. (2015). Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proceedings of the National Academy of Sciences*, *112*(43), 13401–13406. https://doi.org/10.1073/pnas.1503916112

Tibber, M., Anderson, E., Bobin, T., Antonova, E., Seabright, A., Wright, B., Carlin, P., Shergill, S., & Dakin, S. (2013). Visual surround suppression in schizophrenia. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00088

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Yang, E., Tadin, D., Glasser, D. M., Hong, S. W., Blake, R., & Park, S. (2012). Visual Context Processing in Schizophrenia: *Clinical Psychological Science*. https://doi.org/10.1177/2167702612464618

Zhang, Y., Liu, J., Wang, Y., Huang, J., Wei, L., Zhang, B., Wang, W., & Chen, W. (2017). Personality traits and perception of Müller-Lyer illusion in male Chinese military soldiers and university students. *Translational Neuroscience*, *8*(1), 15–20. https://doi.org/10.1515/tnsci-2017-0004