Give me Gestalt! - XPLab 2019

Preregistration Report

Group 04: Cora Burgwinkel, Lina Junctorius, Mirijam Schaaf, Annika Theisen

We aim to replicate the "Give me Gestalt! Preference for Cubist artworks revealing high detectability of objects" experiment by Muth et al. (2012).

Background

Cubism is a highly influential visual art style from the 20th century. The three most famous and most influential painters for this particular style are Picasso, Braque and Gris.
In Cubist artworks, objects are analysed, fragmented and reassembled into an abstracted form. Therefore, Cubist paintings are very open to interpretation.
They are full of everyday objects, but because of the fragmentation, immediate object recognition is difficult.
With this study we aim to show that the subjects' enjoyability of the pictures is linked to the ability of detecting objects, also called Gestalt, in the paintings.

A leading theory behind rating enjoyability and detecting objects is to derive pleasure from searching for and finding recognisable everyday objects. For example, neurologists Ramachandran and Hirstein argue that perceptual grouping processes are in general linked with the neural structures known as the reward system.[1] Reber et al. assert that increased fluency in processing a complex topic enhances appreciation.[2]

[1] V.S. Ramachandran and W. Hirstein, "The science of art. A neurological theory of aesthetic experience," *Journal of Consciousness Studies* 6, No.6-7 (1999) pp.15-51

[2] R. Reber, N. Schwarz, and & P. Winkielman, "Processing fluency and aesthetic pleasure. Is beauty in the perceiver's processing experience?," Personality and Social Psychology Review 8, No. 4 (2004) pp. 364-382.

Research question

The more easily individuals without expertise in cubism can detect objects in the presented paintings from Braque, Gris and Picasso, the higher the likeability is rated.

Design of the Experiment

General remarks about the Design

We will conduct a within subject experiment. The independent variable is the detectability of objects in the painting while the dependent variable is the likeability of the paintings.
In contrast to the original experiment, our replication of the experiment will be web-based. We will be using the architecture babe which is based on JavaScript and HTML while the statistical analysis will be performed with R.

Sampling Plan

Our plan is to at least recruit 20 participants, as done in the original experiment, and even more if we are able to find more subjects. We have a time limit of one week. The participants will be recruited among Cognitive Science students at the University Osnabrueck which are enrolled in the course "Experimental Psychology Lab" and among friends and family. The method we will use is called convenience sampling as a type of non-probability sampling. That means we collect a sample from somewhere convenient to us and that the odds of any member being selected for a sample cannot be calculated. The major advantage of the non-probability sampling technique is that it is very time- and cost-effective. The main disadvantage of this technique is that it is impossible to know how well one is representing the original population, so there may be a shortcoming of representation.[3]
Participants will be recruited via email with a link that provides access to the online experiment. The participants will be no experts in cubist art and will have corrected-to-normal vision as well as normal colour vision.

Procedure

In the original experiment 120 photographs of paintings by Pablo Picasso, George Braque and Juan Gris were used. The stimuli will be limited to 30 images to shorten the duration of the experiment because participants will not get paid and thus, are likely not willing to invest a lot of time in the experiment. As we have the original stimuli, the stimuli set will respectively contain 5 monochrome and 5 coloured paintings of each painter.
Before the experiment begins, we want to make sure that participants meet the required distance to their laptop. The distance should be 55cm. The participants can easily measure this by putting their hand on their monitor and staying an arm length away. We introduce this method to the participants in order to assist them with maintaining the required distance as good as possible. We chose to do so because we are not able to correct their position as this experiment is not conducted in a laboratory setting.

---

[3] Non-Probability Sampling: Definition, Types Online URL:
https://www.statisticshowto.datasciencecentral.com/non-probability-sampling/[last visited 27.06.2019]

The experiment consists of two blocks and in both blocks 30 paintings are shown in a randomised order. Participants are shown written instructions about the task. In the first block the participants are asked to rate the paintings on how much they like them. The participants rate on a 7 point Likert scale from 1 (“not at all”) to 7 (“very much”). During the second block, participants are asked to rate how well they are able to detect objects in the paintings. The participants rate on a 7 point Likert scale from 1 (“very hard”) to 7 (“very easy”) by clicking on one of the seven numbers on the rating. After the two blocks, the corrected-to-normal vision of the participants is tested by a self-constructed vision test and by a short version of the Ishihara color vision test. There is no time limit because the subjects should have enough time to get involved with the paintings and the reaction times do not matter for the statistical analysis. Finally, the participant is asked to indicate how much he is an expert in cubism by again choosing a number on a 7 point Likert scale from 1 (“not at all”) to 7 (“absolutely”).

Here is one example for each of the trials:

How much do you like the painting?

Not at all (1) (2) (3) (4) (5) (6) (7) Very much

How well can you detect objects in the painting?

Very hard (1) (2) (3) (4) (5) (6) (7) Very easy

Measured variables

- likeability, the dependent variable: measured by an ordinal 7-point Likert-scale from 1 (“not at all”) to 7 (“very much”)
- detectability, the independent variable: measured by an ordinal 7-point Likert-scale from 1 (“very hard”) to 7 (“very easy”)

Materials

We will use the original photographs provided by Claudia Muth, the author of the study we are replicating. The pictures were part of the stimuli set used in Muth et al (2012). The photographs have been adapted to 450 pixels width and 600 pixels height. If the proportion was not 4.5:6, the pictures were cropped accordingly. A list of the stimuli can be found below.

- Muth, C., Pepperell, R., Carbon, C. (2012), “Give me Gestalt! Preference for Cubist artworks revealing high detectability of objects” Leonardo, 46

Besides the stimuli we use further materials which are referenced here:

The image of the hand is provided by Clipart.

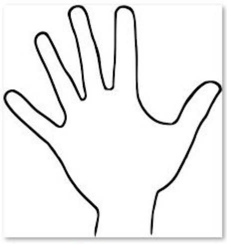
- *Hand Clipart Black and White 2* [picture]. Retrieved from https://www.clipart.email/download/177177.html [last visited 27.06.2019]

The image for the Ishihara Test is provided by specialist and eye surgeon Marek.

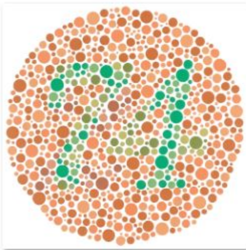
- Eyespecialist and eye surgeon Marek. *Ishihara Test* [image]. Retrieved from https://augenmedizin.at/wp-content/uploads/2014/11/Ischihara-5-300x300.jpg [last visited 27.06.2019]

We created the images for the vision test ourselves. Therefor we created two images with the same width and height as the used stimuli.

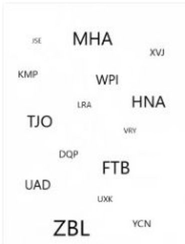
- Junctorius, L., Schaaf, M., Burgwinkel, C., Theisen, A. (2019). *Visiontest and Visiontest 2* [pictures].



Hand.png



ishihara.png



Visiontest.jpg

Visiontest_2.jpg

Stimuli

Total: 30 Stimuli

		Colour Scheme	
		Coloured	Monochrome
Painter	Pablo Picasso	5x	5x
	George Braque	5x	5x
	Juan Gris	5x	5x

List of used Paintings

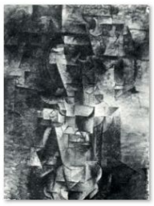
No	Painter	Painting	Year
1	Braque	The old castle at La Roche-Guyon	1909
2	Braque	Still life with mandola, metronome (vase and books)	1909
3	Braque	Woman holding a mandolin	1910
4	Braque	The Rio Tinto factories at L'Estaque	1910
5	Braque	Still life with a bottle	1910
6	Picasso	Still life with liqueur bottle	1909
7	Picasso	The rack (with glass, pipe and letter)	1912
8	Picasso	Bowl of fruit	1910
9	Picasso	Standing nude woman	1910
10	Picasso	Man rowing (with oars)	1910
11	Braque	Factory roofs at L'Estaque	1908
12	Braque	The chateau at La Roche-Guyon	1909
13	Braque	The mandolin (and bottle)	1910
14	Braque	Bottle and glass (on a table)	1911
15	Braque	Still life with harp and violin (glass and ink blotter)	1912
16	Picasso	Man with guitar	1913
17	Picasso	Glass with straws	1911
18	Picasso	Man with clarinet	1911
19	Picasso	Woman with triangular head	1910
20	Picasso	The female student	1911
21	Gris	The pot of geraniums	1915
22	Gris	Still life on a chair	1914
23	Gris	Still Life with Fruit Dish and Mandolin	1919
24	Gris	Water bottle, bottle and fruit dish	1915
25	Gris	The checked tablecloth	1915

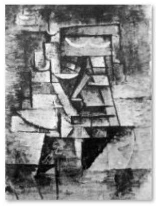
26	Gris	Still life with playing cards	1916
27	Gris	The violin	1916
28	Gris	The checkerboard	1915
29	Gris	Still life with poem	1915
30	Gris	The lamp	1916

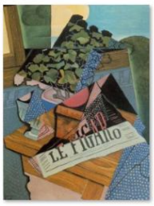

16.jpg


17.jpg


18.jpg


19.jpg


20.jpg



21.jpg


22.jpg


23.jpg


24.jpg

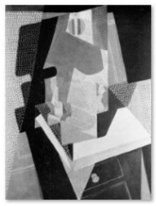

25.jpg


26.jpg


27.jpg


28.jpg


29.jpg


30.jpg


1.jpg


2.jpg


3.jpg


4.jpg


5.jpg


6.jpg


7.jpg


8.jpg

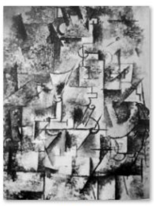

9.jpg


10.jpg


11.jpg


12.jpg


13.jpg


14.jpg


15.jpg

Analysis Plan

Exclusion criteria

The first exclusion criterion is colour blindness because in order to recognise objects in the paintings, it is necessary to have unimpaired colour vision. To test the participants' colour vision, we make use of the Ishihara test. The Ishihara test is a colour perception test developed by Dr. Shinobu Ishihara at the University of Tokyo in 1917. It consists of multiple so called Ishihara plates. As our experiment needs to stay in a reasonable time frame for participants, we decided to only include plate 7. This plate mainly tests if subjects have red-green colour blindness which is one of the most common colour vision deficiencies. The number shown on plate 7 is 74. Viewers may see the numbers 71 or 21 instead. Viewers who see 21 are likely to have dichromacy (two types of cells functioning) or anomalous trichromacy (one of the cones is altered in its spectral sensitivity). Whilst viewers with monochromacy (one type of cone cells functioning) may see nothing. We have decided to exclude people whose answers differ from 74 or 71. We want to include participants that see 71 instead of 74 because given the image's composition and the colour contrast it is likely to see a 1 instead of a 4. If participants see a 71, their colour vision may be slightly impaired but we do not consider it as sufficiently severe as to affect the participants' responses.

After completing the Ishihara test, our participants need to perform a vision test. Therefore, we created an image which has the same width and height as the stimuli and shows different triples of uppercase letters in different sizes distributed on a white background. The different triples contain no similar letter combinations so that they cannot be confounded. The smallest triple has approximately the same size as the smallest object contained in the paintings. The participants are first asked to write down the smallest triple they are able to read in a text field. With this test we want to examine the participants' vision under conditions similar to the trials to make sure that they are able to see all objects appearing in the stimuli. Next, the participants see a similar image in which all triples except the smallest are faded out so that the smallest triple stands out. The participants are asked to write this triple down in a text field as well. With this we want to make sure that they did not miss the smallest triple in the first part of the vision test and again test if they are able to read the smallest triple.

If a participant enters a wrong letter combination in the first and second part of the vision test, his or her data is excluded from the analysis. If a participant enters a wrong letter combination in the first part but the correct combination in the second part, his or her data is not excluded. If a participant enters the correct letter combination in the first part but a wrong combination in the second part, his or her data is excluded depending on the deviation from the correct triple. In this case, if the deviation involves more than one of the three letters, the data is excluded. By doing so, we avoid excluding participants whose vision is good enough for the experiment but that made a typographical error in the second part of the vision test.

Another exclusion criterion is whether the participants are already familiar with or even experts in the cubist art style. We ask them to indicate on a Likert rating scale from 1 (“not at all”) to 7 (“absolutely”), if they are an expert in Cubism. When having answered with a 6 or 7, the participant is excluded. The method of self-reporting used here is appropriate because there is no reason for the participants to not answer this question honestly. The participant knows best if he or she has knowledge about the topic.

Finally, we will exclude participants who need less than 5 minutes to complete the experiment because we think that you cannot work through it seriously in such a short time frame. Additionally we will exclude participants who rate 50 trials and more with the same number because this leads us to believe that the participant just clicked through the experiment without engaging with the paintings.
However, we will not exclude participants who do not use the most extreme values (1 and 7) because often one avoids the extremes and might not feel that strongly about the paintings.

Confirmatory hypothesis testing

For our main analysis, the hypothesis “The more easily individuals without expertise in cubism can detect objects in the presented paintings from Braque, Gris and Picasso, the higher the likeability is rated.” will be tested by using a Bayesian regression model. In the code sample below, it can be seen that we utilise the brms package to fit the models. We will analyse one main model to examine our hypothesis. It is a hierarchical model with only fixed effects treating the object-ratings as interval-scale. We want to regress the dependent variable likeability against the independent variable detectability of objects (formula = likeability ~ objects). Additionally, a couple of hierarchical models will analyse more specific fixed and random effects. For example, one hierarchical model will test whether there is a correlation between the likeability of the paintings and the artist.

Besides the Bayesian analysis, we will also use the frequentist approach from the original study for our analysis. This is only done to compare it to our main analysis and in order to stick to the original study. The authors tested the hypothesis by using a Pearson correlation coefficient.
As done by Muth et al (2012), we will report the value of r and the p-value indicating whether the correlation is significant or not. The results will also be depicted in a regression graph: on the x-axis, the detectability of objects will be shown and on the y-axis, it will be shown how much subjects liked the painting. One point in the graph shows the values for a single painting. We will further report the amount of explained variance r^2.

We did not choose this approach as our main analysis because it has a lot of limitations and is not totally appropriate for the experiment. The Pearson correlation presupposes interval data for both variables. When working with a Likert scale, the data is actually only ordinal. Furthermore, the authors did not test the assumption of a linear relationship as well as a bivariate normal distribution. Especially the last assumption requires to also execute a power analysis in order to determine the sample size needed to find a certain effect. All of this was not done in the original study. We will stick to the original analysis just to compare our results, but rely on the Bayesian approach for our main analysis.

First draft of code using the statistics software R

The Bayesian model in R:

```
library(tidyverse)
library(brms)
library(ggplot2)

d = read_csv("test_results_1_adjusted.csv") %>%
  filter(trial_name %in% c("rating_scale_object", "rating_scale_like")) %>%
  select(submission_id, trial_name, response, picture_nr, artist)

# cast data into appropriate type
d_wide = spread(d, key = trial_name, value = response) %>%
  mutate(likeability = factor(rating_scale_like, ordered = T),
         objects = factor(paste0("C", rating_scale_object)),
         objects_forward = objects)

# inspect data
ggplot(d, aes(x = objects, y= likeability)) + geom_point() + geom_smooth(method = "lm")

# define a forward contrast : this way each next level (number) will be compared to the
# previous level (number)
forward_contrasts = matrix(c(0,0,0,0,0,0,
                             1,0,0,0,0,0,
                             1,1,0,0,0,0,
                             1,1,1,0,0,0,
                             1,1,1,1,0,0,
                             1,1,1,1,1,0,
                             1,1,1,1,1,1
), byrow = T, ncol = 6)
colnames(forward_contrasts) = paste0(".", 2:7)
rownames(forward_contrasts) = paste0(".", 1:7)

contrasts(d_wide$objects_forward) = forward_contrasts

# 1. simple analysis:
m = brm(data = d_wide,
        # hierarchical model with only fixed effects (treating object-ratings as interval-scale / metric)
        formula = likeability ~ as.double(objects),
        # hierarchical model with by-item (pictures) and by-subject random intercepts
        formula = likeability ~ objects + (1 | picture_nr) + (1 | submission_id),
        # hierarchical model with by-subject random intercepts and fixed effect of artist
        formula = likeability ~ objects + artist + (1 | submission_id),
        family=cumulative("logit")
)

# 2. more complex analysis with forward contrasts
m_forward = brm(data = d_wide,
                # hierarchical model with only fixed effects
                formula = likeability ~ objects_forward,
                # hierarchical model with by-item (pictures) and by-subject random intercepts
                formula = likeability ~ objects + (1 | picture_nr) + (1 | submission_id),
                # hierarchical model with by-subject random intercepts and fixed effect of artist
                formula = likeability ~ objects + artist + (1 | submission_id),
                family=cumulative("logit")
                )
```

The Pearson correlation in R:

```{r}


data <- read_csv("/Users/Mirijam/Documents/BachelorCognitiveScience/4.Semester/XPLab/Final_Project/GiveMeGestalt_Exp/test_results.csv")
view(data)
```

#load libraries
```{r}
library(tidyverse)
library(ggpubr)
```

#data formatting
```{r}
data_temp <- as_tibble(data) %>% mutate(response = as.integer(response))
data_temp

x <- filter(data_temp, trial_name == 'rating_scale_like') %>%
  select(c('response', 'picture_nr'))%>%
  group_by(picture_nr) %>%
  summarise(response = mean(response))
x
y <- filter(data_temp, trial_name == 'rating_scale_object') %>%
  select(c('response', 'picture_nr')) %>%
  group_by(picture_nr) %>%
  summarise(response = mean(response))
y

data_formatted <- merge(x,y, by = 'picture_nr')
data_formatted
```

#test variables for normal distribution (assumption of the Pearson correlation)
```{r}
shapiro.test(data_formatted$response.x)
shapiro.test(data_formatted$response.y)
```

#correlation test and regression graph:
```{r}
ggscatter(data_formatted, x = "response.x", y = "response.y",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "Detectability", ylab = "Liking")

res <- cor.test(data_formatted$response.x, data_formatted$response.y,
                    method = "pearson")
res

#extract the p.value
res$p.value

#extract the correlation coefficient
res$estimate

#the amount of variance explained
res$estimate^2

```