

# Curate an art gallery exhibition

## Instructions:

You have been invited to curate an art exhibition in a major museum. 50 different art works have been submitted but only 5 can be selected

Your task is to look through all the art works carefully and **select 5** for the exhibition.

Once you have selected the 5 art works, you will be asked to provide, for each one

a. A **title** that describes the art work

b. A **short description**, to be presented next to the work as as a caption (up to 20 words).

**Part 1:** View through all 50 art works.

**Part 2:** Select 5 art works for the exhibition. Thumbnails of all art works will be presented.

**Part 3:** Provide a title and a short description for each selected art work.



START

**Figure A1.** Instructions for the Gallery Curation Task.

## Appendix A. The Aesthetic Evaluation Tasks

Experiment 2 (gallery curation) and experiment 3 (ranking task) were designed to increase viewers' engagement. MTurk participants were provided with the instructions in Fig. [A1](#) for the gallery curation task and Fig. [A2](#) for the ranking task. Additionally, participants were asked to provide an explanation for their ranking decisions in the ranking task on five of the 30 trials, selected at random (see Fig. [A3](#)).

## Appendix B. Results of Experiment 1: Likert-scale Ratings

Here we report the complete statistics of Experiment 1. Tab. [B1](#) shows the averaged ratings of each image category as well as their standard deviations. The average ratings between each pair of categories are compared in Tab. [B2](#), using the Mann-Whitney U rank test. We report both the U statistic and the p value. Similarly, the average statistics and comparisons between categories in response times are reported in Tab. [B3](#) and Tab. [B4](#).

# Image Rank

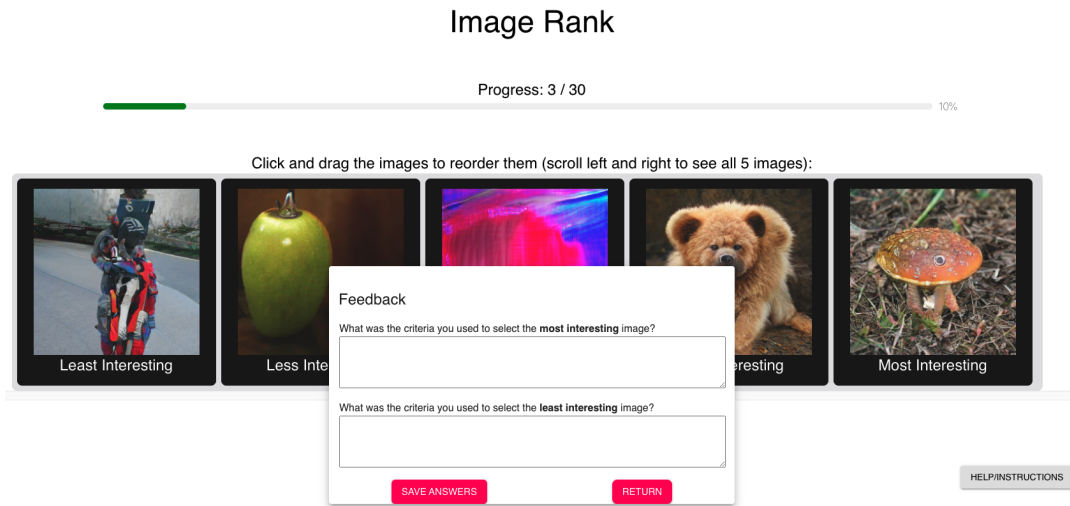
Rank the images according to how interesting they are!

How To:

You will be shown a set of 5 images. Your task is to rank the images from most to least interesting. You will need to:  
Order 5 images by how interesting each of them is to you. **Drag** the images around to order them!

START

**Figure A2.** Instructions for the Ranking Task.



**Figure A3.** Participants are asked to provide an explanation for their ranking decisions.

Category	Interestingness		Powerfulness		Engagement	
	Mean	std	Mean	std	Mean	std
Recognizable	3.36	0.60	2.96	0.52	2.98	0.57
Dichotomous	2.87	0.40	2.76	0.37	3.07	0.46
Indeterminate	2.99	0.39	3.13	0.36	2.98	0.46
Abstract	2.88	0.39	3.01	0.36	2.94	0.36
AbstractFlat	2.61	0.56	2.57	0.58	2.47	0.54

**Table B1.** Average ratings of Experiment 1. Experiment 1 asked participants to rate images based on how “interesting”, “powerful” and “engaging” they are. The mean and standard deviations (std) of the ratings for each image category are reported here.

	Interestingness		Powerfulness		Engagement	
	U	p	U	p	U	p
Recognizable vs. Dichotomous	659.5	0.0020	547.0	0.15	423.0	0.69
Recognizable vs. Indeterminate	610.5	0.018	340.0	0.11	450.0	0.99
Recognizable vs. Abstract	645.5	0.0039	386.5	0.35	492.0	0.54
Recognizable vs. AbstractFlat	717.5	$\ll 0.001$	590.5	0.038	675.5	0.00087
Dichotomous vs. Indeterminate	387.0	0.35	203.5	0.00027	482.0	0.64
Dichotomous vs. Abstract	426.5	0.73	268.5	0.0074	532.0	0.23
Dichotomous vs. AbstractFlat	563.0	0.096	498.0	0.48	723.5	$\ll 0.001$
Indeterminate vs. Abstract	498.0	0.48	517.5	0.32	469.0	0.78
Indeterminate vs. AbstractFlat	610.5	0.18	704.0	0.00018	672.5	0.0010
Abstract vs. AbstractFlat	571.5	0.073	645.0	0.0040	685.0	0.00052

**Table B2. Comparisons of the average ratings between different image categories.** We use the Mann-Whitney U rank two-sided test and report the U statistic and the associated p-value for each comparison, corresponding to the values reported in Table [B1](#)

Category	Interestingness		Powerfulness		Engagement	
	Mean	std	Mean	std	Mean	std
Recognizable	4.68	1.68	3.52	0.522	3.93	1.21
Dichotomous	4.57	1.01	5.17	1.75	4.27	1.22
Indeterminate	4.56	0.740	5.36	3.61	3.90	0.986
Abstract	4.93	2.67	3.90	0.489	4.11	0.567
AbstractFlat	5.08	5.58	3.47	0.547	4.07	2.62

**Table B3. Average response times in Experiment 1.** The mean and standard deviations (std) of the averaged response times (in seconds) for each image category are reported here.

	Interestingness		Powerfulness		Engagement	
	U	p	U	p	U	p
Recognizable vs. Dichotomous	422.0	0.68	74.0	$\ll 0.001$	341.0	0.11
Recognizable vs. Indeterminate	387.0	0.36	97.0	$\ll 0.001$	429.0	0.76
Recognizable vs. Abstract	447.0	0.97	254.0	0.0038	296.0	0.023
Recognizable vs. AbstractFlat	657.0	0.0023	484.0	0.62	483.0	0.63
Dichotomous vs. Indeterminate	422.0	0.68	494.0	0.52	549.0	0.15
Dichotomous vs. Abstract	481.0	0.65	742.0	$\ll 0.001$	398.0	0.45
Dichotomous vs. AbstractFlat	674.0	0.00095	827.0	$\ll 0.001$	615.0	0.015
Indeterminate vs. Abstract	511.0	0.37	692.0	0.00036	308.0	0.036
Indeterminate vs. AbstractFlat	703.0	0.00019	801.0	$\ll 0.001$	501.0	0.046
Abstract vs. AbstractFlat	670.0	0.0012	660.0	0.0020	664.0	0.0016

**Table B4. Comparisons of the average response times between different image categories.** We use the Mann-Whitney U rank two-sided test and report the U statistic and the associated p-value for each comparison, corresponding to the values reported in Table [B3](#).

	Mean	std
Recognizable	3.95	0.47
Dichotomous	3.83	0.35
Indeterminate	3.80	0.35
Abstract	3.88	0.33
AbstractFlat	3.83	0.37

**Table C1. Average response times in Experiment 2.** The mean and standard deviations (std) of the averaged response times (in seconds) for each image category are reported here.

## Appendix C. Results of Experiment 2: Gallery Curation

The average viewing time of each image category is reported in Tab. [C1](#). Compared to Tab. [B3](#) there is no significant difference according to two-tailed t-tests.