

Lab 07
Social Psychology
Assigned – now
Due – TBD

In this lab, we will collect images from the internet of people (men and women) crudely categorized as “good” and “bad”. The class will then assess, based solely on their images, some basic personality attributes. Details are provided below, but each image should be a front-facing, level image of the person with a neutral expression. The following is an ideal example:



After collection, the images will be cropped and otherwise processed to remove distractions, so background, image centering, etc. in the original are of no great concern. What is important is the head-on posture and neutral expression – we will later try to extract the facial features that correspond to the personality results, so position and expression must be carefully controlled.

Images should all be of European ancestry or “white”. This is to minimize racial variation, which will contribute to the “error” in our statistical analysis. Furthermore, images should be of people that are obviously neither old or young. So, they should be from, say, 20 to 40 years old.

- 1) Each of you should collect from the internet two male and two female images based on some search criterion that might indicate “good”. One might, for instance, look for award winners or scientists or teachers giving the individuals the benefit of the doubt as to the actual “goodness”. Next, find two males and two females using some criteria you can reasonably associate with “bad” - murder, criminal, etc. Select the images based solely in their quality (position and expression) and try to make the same sex in the good/bad category of approximately similar ages. Name your images according to the provided .csv file. (The file can be read directly into R, but if you want to import it into a spreadsheet, you may need to specify the .csv file is encoded in UNICODE (UTF-8) so the spreadsheet will know how to parse the characters).

Make sure you do not use photos used by another person. If two people select the same photo one has to find another. If both photos are taken you have to find another for one and let the other person replace theirs. To help avoid overlap, you should send your pics to the TA, who

can show them on the screen as they are collected.

- 2) Using GIMP software, cut out the face as a circle making sure to get the ears and chin clearly represented while removing as much of the background and hairstyle as possible. Save these results. Convert the image to grayscale and scale to at least 480x480. Follow the example of your last lab to produce for each subject something like:



- 3) Combine the edited images into a class data set assigning each a random number and recording the sex and good/bad category of each.

In the end, there should be up to:

22 unique “good” female photos

22 unique “bad” female photos

22 unique “good” male photos

22 unique “bad” male photos

- 4) The TA will then (probably and hopefully on Wednesday) show the images on the projector in random order for approximately 5 seconds each. For each, on a five-level scale indicate your answer to the following question related to each image noting the image file name:

“I would be willing to loan this individual $\frac{1}{2}$ my current savings for one week.”

1. Definitely not
2. Probably not
3. Maybe, maybe not
4. Probably
5. Definitely

For the images you chose, mark NA. E.g.,

123_F.png 3

456_M.png NA

- 5) In a different class (Friday?) and using a different random presentation of images, do the same for the following statement:

“I would vote for this person to be the leader of our lab group of five students.”

1. Definitely not
2. Probably not
3. Don't know
4. Probably
5. Definitely

Again, for the images you chose, mark NA.

- 6) Combine the scores making sure to keep intact the image id and rater data. Note that a more sophisticated way to organized these data might be an array in which one matrix or data.frame corresponds to each raters score for an attribute for each image, with “depth” associated with a variable. Something like: `data[i,j,k]`, where `i=image`, `j=rater`, `k=variable`. We will, instead, use a simpler structure and just organize the data into two sets – one for each question above. Thus, these will be a data frame like `data_4[i,j]` and `data_5[i,j]`, where `i=the same image in both sets` and `j=the same rater in both sets`. The value at `i,j` is the rating for that image by that rater on the associated question.

That is, the data should look for each question something like:

Image_id	Rater01	Rater02	Rater03	...	Rater11
329	3	2	1	...	2
...					

- 7) The analysis of the above data can get quite complicated, but we will simplify it. For each image, compute the mean score across all raters, and combine with `image_id`, `sex`, and `category` (good/bad) in a single data set. Save it.
- 8) With the data set above, test if there are significant differences between “good” and “bad” class individuals in the responses to the two questions. Let us call the first question “trustworthiness” and the second “leadership”. Are there differences between in perceived “trustworthiness” between “good” and “bad” men? Women? Leadership?
- 9) Are there perceived differences between “good” men and “good” women with respect to perceived “trustworthiness”? Leadership? Sex differences between then “Bad” class?