**Preregistration**

**1. Study information**

*Title (required)*

How does binary thinking influence face adaptation to gender?

*Authors (required)*

*Description (optional)*

In encounters with others, people use social information to ascribe social categories. (Bodenhausen et al., 2012). The process and consequence of social categorization are frequently studiedd and the literature is dominated by methods which assume gender to be binary., Recently The binary concept of gender has been challenged both by queer activists (e.g. references) and researchers(e.g. Morgenroth et al., 2020; Hyde et al., 2018). These criticisms say that gender is socially constructed and continually re-created in social interactions (e.g., Butler…). For example, when researchers measure perceptions of gender of faces, the design of a study and how response option are used, re-create and communicate ideas about gender. The purpose of this study is to look deeper into how various ways of constructing gender, specficially through asking questions, can influence categorizations of gender.

*Hypotheses (required)*

The study exploratory and descriptive. As such, no specific testable hypotheses are made. The aim of the study is instead to carry out a broad exploration of the pattern of gender categorization when participants are given a wide range of response options with which to categorize gender.

In lieu of statistical tests of hypotheses, the data will be summarized and presented visually. Statistical modelling will be carried out, primarily in order to estimate the uncertainty of our measures. We do expect to see a broad trend where participants given the option to make gender categorizations beyond the binary choose to do so. Furthermore, we expect that participants who are given the option to categorize faces beyond the binary will have visibly lower categorical perception curves.

## **Design Plan**

*Study design\**

The experiment investigates gender categorization using a number of outcome measures. Across two studies, participants will be shown either faces or words and asked to categorize each in terms of gender. The five experimental conditions A number of different types of response options will be given (see measured variables). The various outcome measures are varied on a between-subjects basis, participants are randomly assigned into one of the five conditions.

*Stimuli*

In study 1, participants are shown 140 faces. These consist of morphs of women and men. Ten 7-step morphs will be made from woman-man face pairs drawn from the London face database (ref).

In study 2, participants are show 12 sets of words. Each set of word is a an adjective stereotypically associated with either the communion or agency traits. The words are randomly selected from a pre-validated list of descriptors with the following restrictions. Two sets contain only communal descripters, two contain five communal and one agentic descriptor, two contain four communal and one agentic descpiptor, two contain three communal and three agentic and so on.

## **Sampling Plan**

**Check**Registration prior to creation of data

*Data collection procedures\**

**Study 1**

Study one will a laboratory experiment where participants are shown a series of faces and asked to rate them in terms of gender. Participants are randomly assigned one of three response options conditions. Based on available funding, 100 participants will be tested for study 1.

**Study 2**

Study one will an online experiment where participants are shown a series of descriptions and asked to rate them in terms of gender. Participants are randomly assigned one of three response options conditions. Based on available funding, XX participants will be tested for study 2.

## **Variables**

*Manipulated variables*

**Study 1**

The main manipulated variable will be the response options used to measure gender categorization/rating of faces. The following response options will be used.

-  *Binary categories*. Participants categorize gender as two discrete categories consisting of the options “woman” and “man”.

- *Multiple categories*. Participants categorize gender as four discrete categories, consisting of: “Woman”, “Man”, “Other” and “I don’t know”.

- *Binary scale*: Participants categorize gender using a slider, a single line where they indicate how they perceive the face on a continuous scale from “Clearly a woman” to “Clearly a man”.

- *Multiple scales*. Participants categorize gender using a slider, similar to the binary scale, with the sole difference being the scale ranges from “Not a woman/man” to “Clearly a woman/man”. Each stimulus is categorized twice, once for each gender identity.

- *Free text.* Participants are given an open text and are free to write in whichever categories they find most appropriate.

**Study 2**

Study 2 will contain identical conditions as study 1. However, the main difference is that the categorization/rating will be of descriptors rather than faces.

*Measured variables*

**Study 1**

The main measured variables are participants categorizations/ratings of faces. Because study 1 uses five different scales, there will be six outcomes, one for each condition, except for the *Multiple scales* condition, which has two outcomes.

**Study 2**

Again, study 2 uses an identical outcome measure, to a different stimuli.

**Analysis Plan**

*Statistical models\**

**Study 1**

Due to exploratory nature of the study, the primary purpose of modeling is to estimate the uncertainty of the measures, rather than to carry out inferential testing. Because the various measures consist of radically different outcome, the same modelling strategy will be applied to each measure. Two values are of interest. First, the overall tendency to which the various categories are used. Second, how often the categories are used at various morph levels. The modelling strategy reflects these goals, where the main variables of interestest are both overall intercepts and unique intercepts at each level of morph.

The goal of modeling is to estimate these parameters and the uncertainty of these parameters. Consequently, no inferential tests are specified, nor are any specific comparisons made. This modeling strategy consists of several multilevel models of increasing complexity. In the simplest model, unique intercepts are calculated for each morph level. Additionally unique intercepts are calculated for each subject as well for each unique face, with hierarchical priors on both. A somewhat more complicated model allows the effect of morph level to vary by race. A last, most complicated model allows the effect of morph level to vary by subject. These three models will be compared using leave-one-out cross validation. The purpose of fitting multiple models and carrying out model comparison is to account for the uncertainty in which model is correctly specified.

**Study 2**

The analyses for study 2 is identical with the exception

*Transformations*

*Inference criteria*

*Data exclusion*

Study 2

*Exploratory analysis*

The effect of race