# rPPG data collection

## Hardware

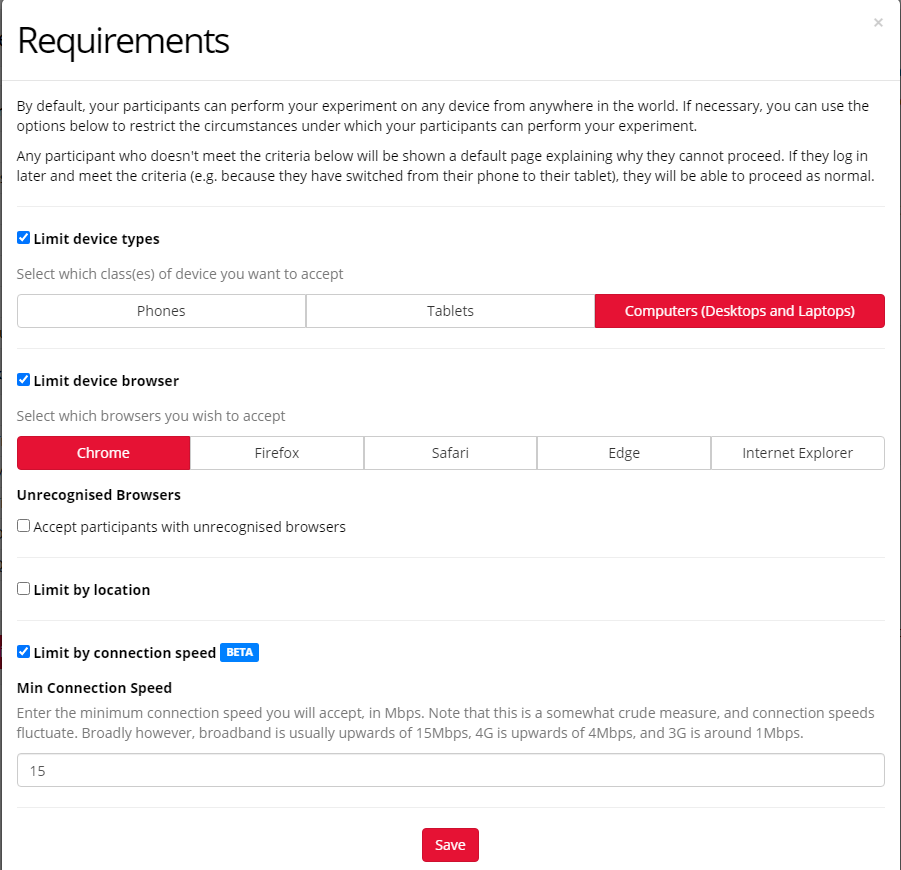
In order to take part in the experiment, participants will need:

* Desktop or laptop computer
* Webcam
* Google Chrome Browser
* High speed internet connection

In Gorilla, go to Recruitment:



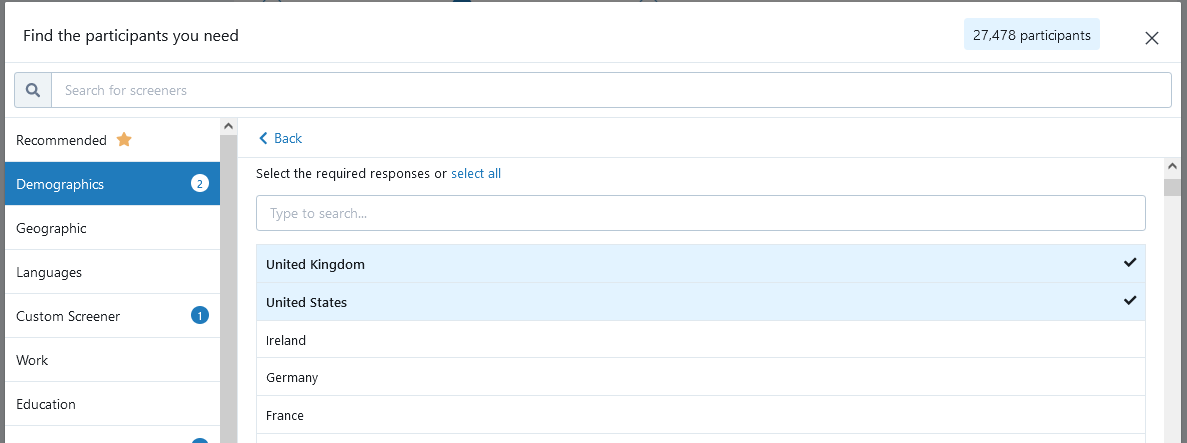
Then select:



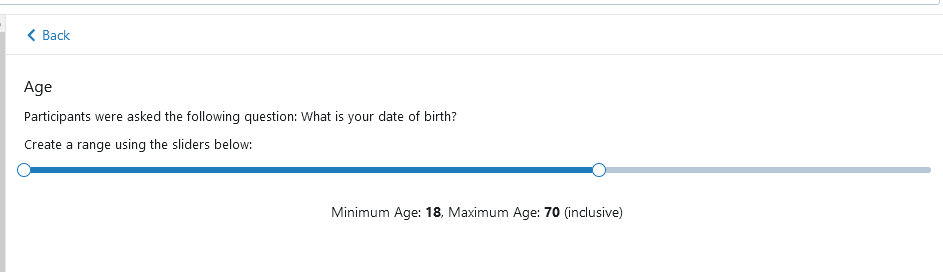
## Data collection

### Inclusion criteria

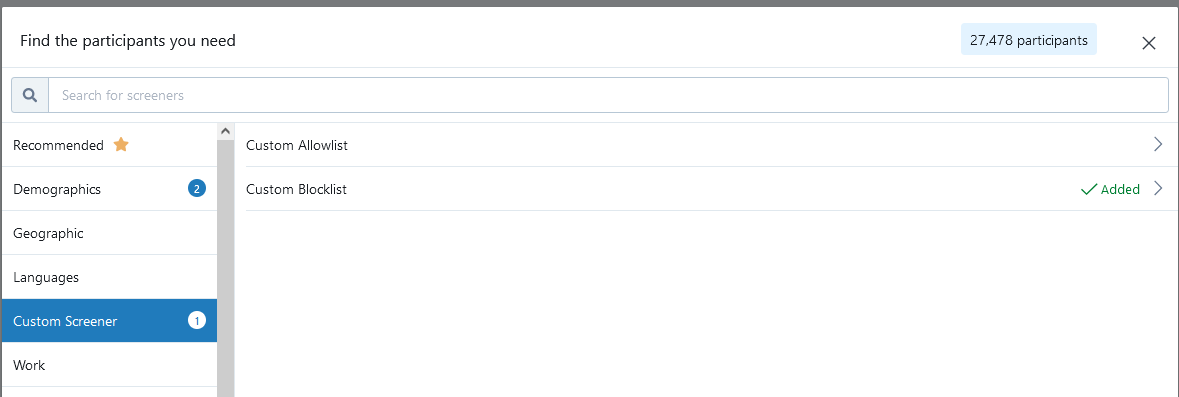
Country of residence:



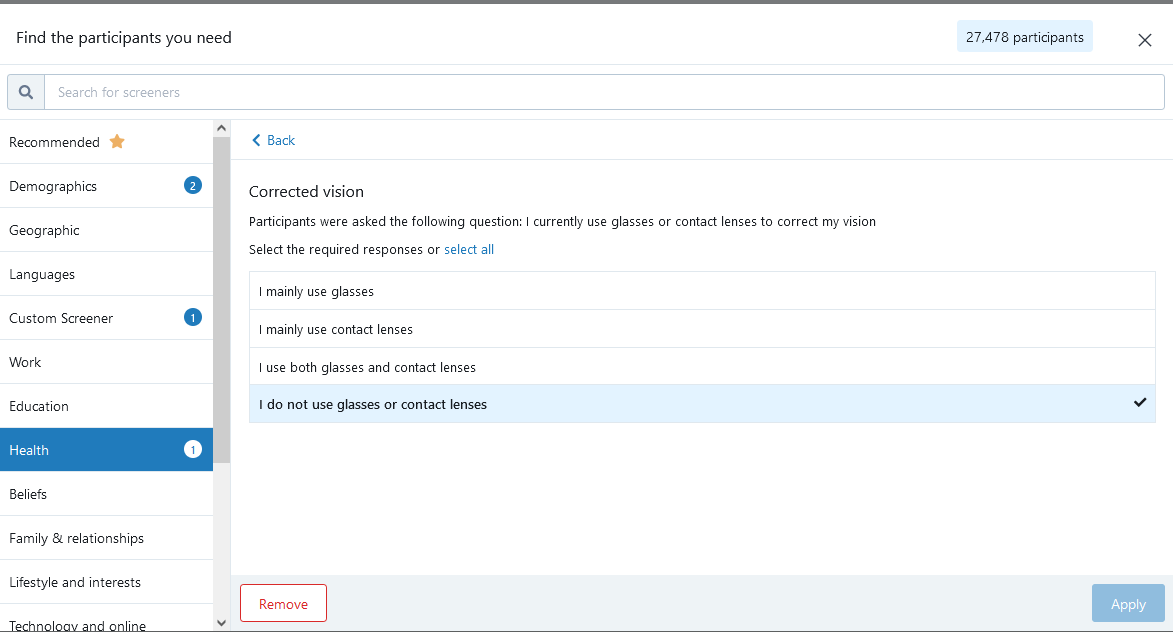
Age



Custom Blocklist: don’t allow participants that we know are bad from past experiments or, eventually, that have already participated in a previous version of the experiment.

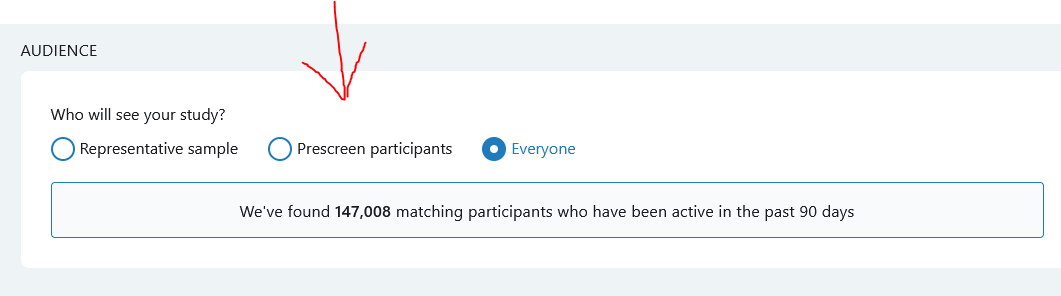


Health -> corrected vision:

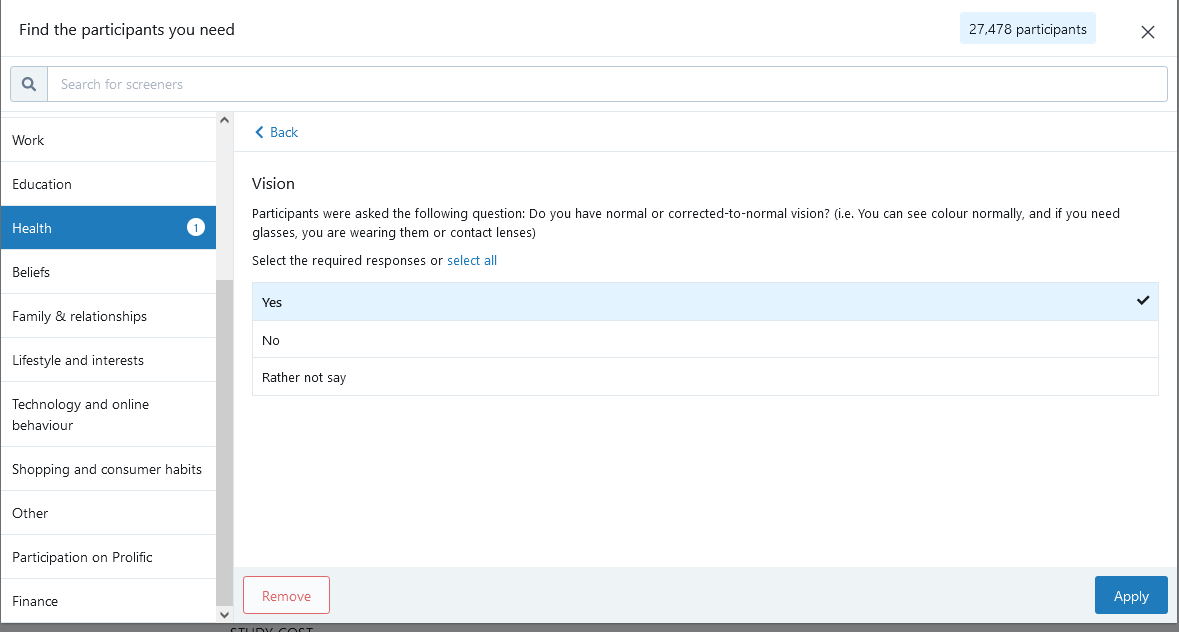


* Normal vision or corrected with contact lenses: participants cannot wear glasses during the experiment

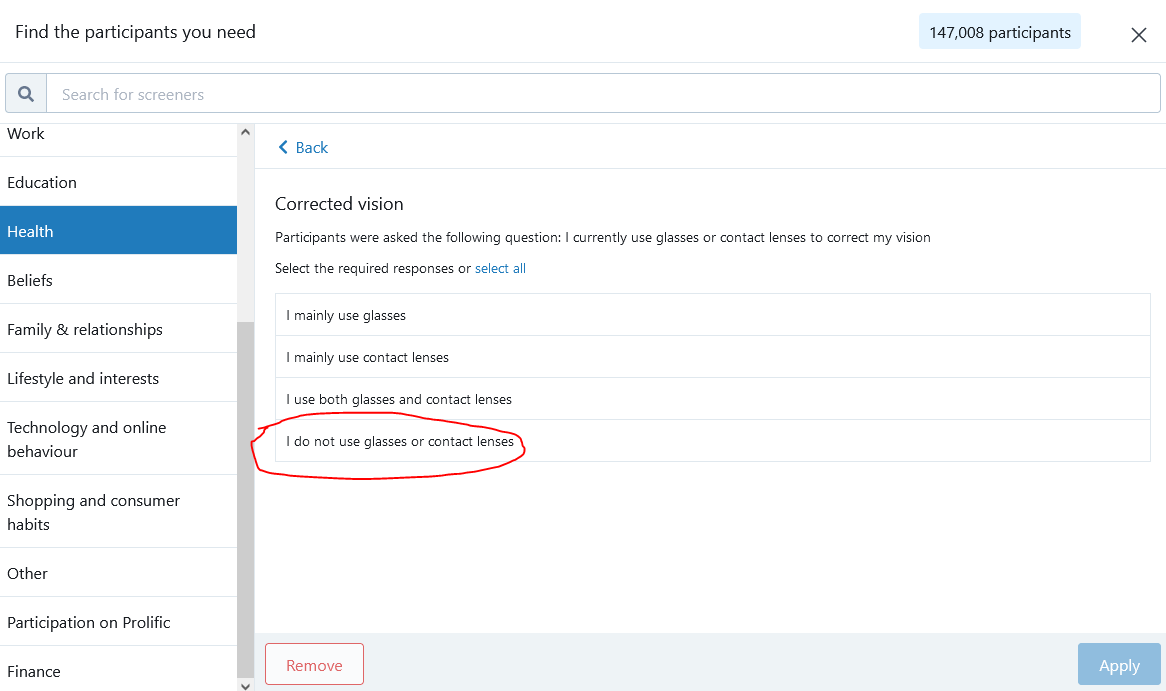
In prolific you can select participants with normal vision.  
Select: Prescreen Participants



Scroll to find “Health”, Scroll all the way down to “Vision”, Select “Yes”



Then, in corrected vision select: I do not use glasses or contact lenses:



To sum up, it should look something like this:  


### Instructions

These guidelines must be followed to maximize video quality:

* Recording with daylight
* Room well illuminated with natural light
* Participants must sit as still as possible
* Face must not be covered by shadows
* Face the camera directly
* Sit close to the camera
* Computer must be on a table: NOT KNEES OR UNSTABLE SURFACE
* No part of the face can’t be covered (e.g., by wearing a mask, by hair or touching the face)

As a reference, see following image:  


ADDITIONAL WARNING

1. As I said, the room must be well illuminated, HOWEVER, the videos should also not be overexposed.
2. The light of the screen is very detrimental for the rPPG. Some videos look decently illuminated, but the light is actually mostly coming from the screen. These videos cannot be trusted. The light from the screen flickers at a certain frequency, depending on the screen, and this will greatly affect the rPPG performance.

# After data collection

## Exclusion criteria

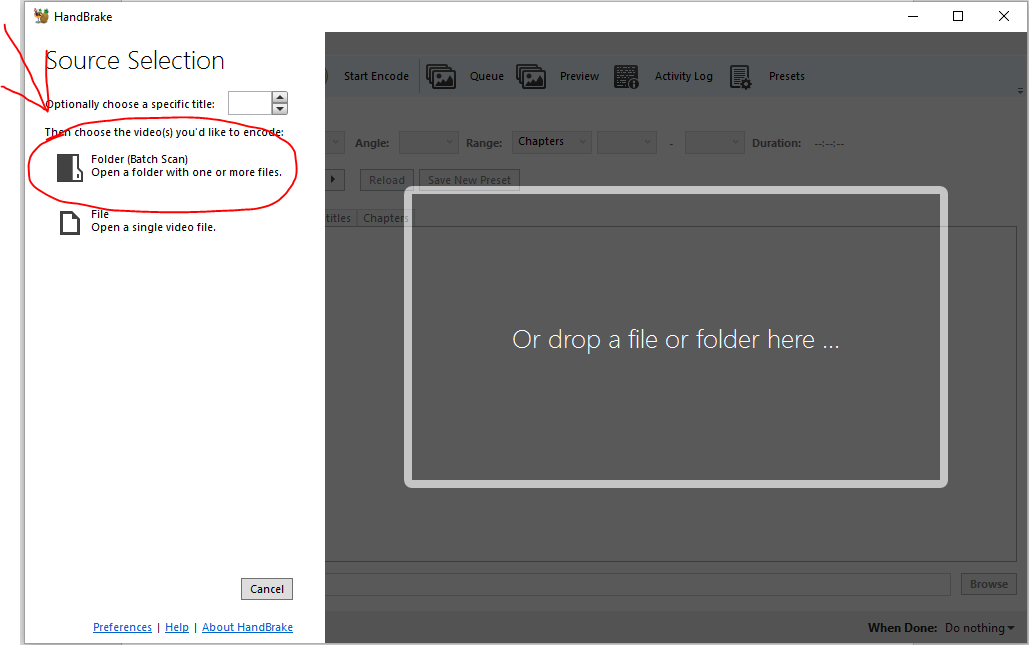
* Visually inspect the videos and exclude participants that:
  + Perform the experiment with a laptop on their legs.
  + Have the webcam off
  + Wear a face mask
  + Are not in front of the camera
  + Are in a dark room
  + Wear glasses

# rPPG video processing

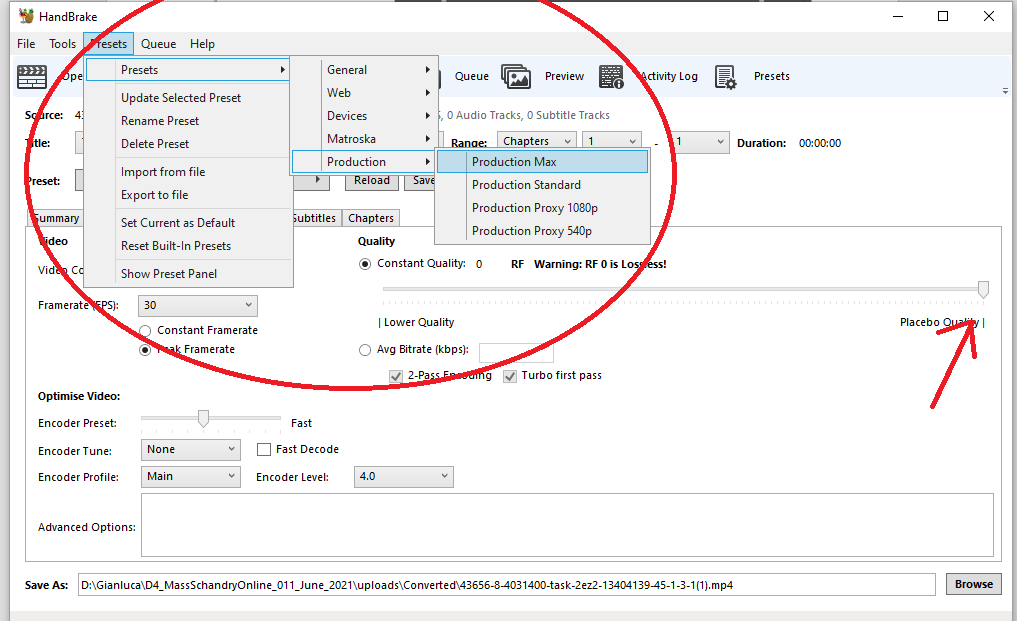
## Handbrake

As a first step, you need to convert the videos from .webm to .mp4.

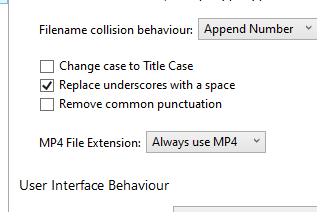
Open HandBrake, select “Folder Batch Scan”. This will open a GUI that will allow you to select the folder containing your .webm videos.



After selecting the folder, select: presets -> Production -> Production Max. Then, move the “quality” dial to the maximum (Placebo).



Make sure that the output file is .mp4 by default.



Under Queue, select add all to queue and then run the queue

## rPPG

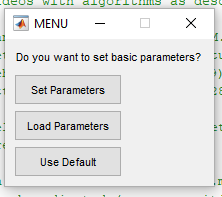
Open the script called RunMe\_MN\_GF-11\_June\_2021

If you are using your personal folder on the computer in 221, this file will have your own name.



### Set Parameters

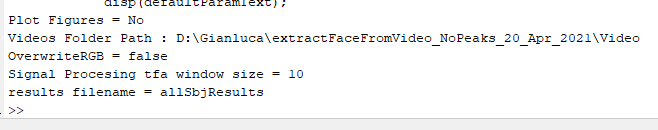
Once you click RUN the script will ask you a series of questions. First, it will ask you if you want to set basic parameters.



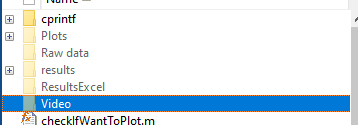
**If this is the first time you run the script on a given set of data, I recommend clicking “Set Parameters”.**

If you have already set the parameters and saved a file “settings” (see later), you can click “**Load Parameters”**. In this case, the script will load the previously saved settings and will start running without asking any further question).

If you set “**Use Default**”, the script will select the following parameters:



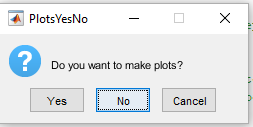
Notice that in this case you WILL HAVE TO put your files in this subfolder:



### Make Plots

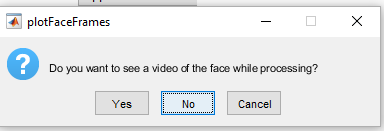
If you have clicked on “Set Parameters”, the script will ask you if you want to make plots. If you click **“Yes”,** obviously the script will show you the plots for each participant. The figures will be closed when moving to the next participants. All the plots will be saved in a folder that you select later.

If you select **“No”,** no figure will be displayed, **however the script will still save the figures, thereby saving processing time**. You can still look at them later, so, I prefer selecting “No”.



### Plot Face Video Live

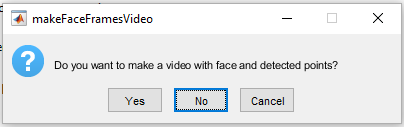
The script will ask you if you want to see a video of the face while processing. This refers to a frame by frame video of the face showing the contour points identified by the computer vision algorithm. Plotting this video is extremely time consuming. Unless you have particular needs, I would select **“No”**.



You can still look at the video later if you want to by selecting **“Yes”** to the next question.

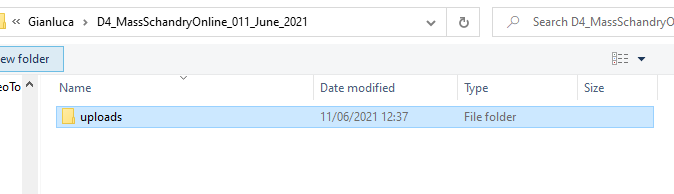
### Save Face Video

This question asks whether you want to save the video of the face with the points, even if you are not displaying it live. I would recommend saving it so that you can take a look at it later. HOWEVER, taken together, these videos occupy a lot of space on the disk. Therefore, once you have looked at them and made your mind regarding the quality of the video, I would delete these files.



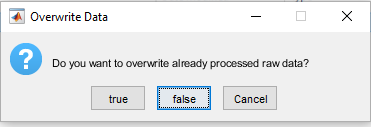
### Select Video Folder

At this point, a GUI appears that allows you to select the folder containing the .mp4 files to process. Navigate to the desired folder.



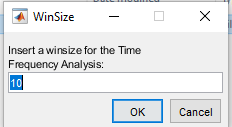
### Overwrite Raw Data

The next step asks you if you want to overwrite the raw data. As you’ll notice, the script crashes on some files that cannot be processed. Once you restart, the script will go back to file one. If you select **“true”**, every time you restart the script it will reprocess all the videos, which will be extremely time consuming. Unless you have some good reasons, I would recommend selecting **“false”**. This way, the script will check if there is already a matlab file for that participant (you’ll find these in the subfolder called “Raw data” . If there is already such a file, the script will load the data for that participant and move on, without reprocessing it.



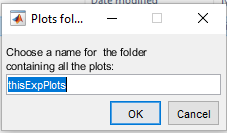
### Select WinSize

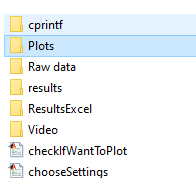
Here you can select the winsize for the Time Frequency Analysis. By default, this is set to 10. If the videos are short, for instance 10 seconds, the script will crash. In this case, try selecting a winsize = 5, otherwise keep the default value.

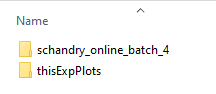


### Select Plots Folder

You can select a subfolder that will contain all the plots that you are going to analyse. This folder will be created inside the folder called “Plots”. For instance, if I call it “thisExpPlots”



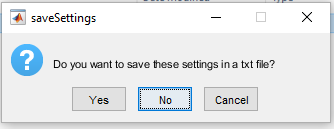
You will find this folder in Plots: 



If you open this folder you will find a video with the face for each participant and a subfolder for each participant containing their respective plots.

### Save Settings

Finally, the script will ask you if you want to save the settings in a txt file.   
If you select **“Yes”**, everything you have just selected will be saved in a txt file called “scriptSettings”. The next time you run the script, you can just “load settings” and enjoy life.



### Start Processing

At this point, the script will start processing the data…



# Data processing exclusion criteria

Exclude all videos with Heart Rate outside this range: 50 – 120

Exclude all videos with Frame Rate < 20

HR measures:

You have several measures, prominently HR Power Based and Smoothy. Our validation experiments show that smoothy works better on Gorilla videos. Using that measure is recommended.

Inspect heart rate range between the 3 videos for each participant.   
If the heart rate between the 3 videos varies more than 10 points, check the randomization for that participant.   
  
Example case

|  |  |
| --- | --- |
| 25 seconds | 80 |
| 35 seconds | 70 |
| 45 seconds | 75 |

Check randomization

Example 1

|  |  |
| --- | --- |
| 35 seconds | 70 |
| 45 seconds | 75 |
| 25 seconds | 80 |

Even if the difference between 25s and 35s is 10 points, they were presented farther apart in time. It seems more likely that the HR changed (participants might have stood up and moved??).

Example 2

|  |  |
| --- | --- |
| 45 seconds | 75 |
| 25 seconds | 80 |
| 35 seconds | 75 |

In this case, 25s and 35s are consecutive. The HR changes by 10 points in one trial. Seems unlikely: consider for exclusion.

**Peaks sanity check**

This is an additional feature that I added in an attempt to identify bad videos. I need to double check with Marnix if this makes sense, hence, use this with caution.

In plots, you’ll notice I have added this plot:  
Chart, histogram

Description automatically generated

Here you can see the power spectrum (as in another plot previously available) with some additional info.

You can notice that there are some red dotted vertical lines and some horizontal lines. One of the vertical lines indicates the frequency with the highest power.

The horizontal lines divide the highest peak in 4 equal parts. The script checks if there are other frequencies with a power higher than ¾ of the highest peak/frequency.

In the example displayed above, you can see that frequencies meet this criterion. These frequencies are also indicated with vertical dotted lines.

Because there are so many frequencies similar to the highest one (which corresponds to our estimated HR), this video gets flagged in the excel file.

Indeed, in the excel I have added 3 columns: flag, nPeaksAboveThreshold, maxFreqDiff

A picture containing table

Description automatically generated

**nPeaksAboveThreshold** tells you how many frequencies have a power higher than 3/4 of the highest power (e.g. the first video has a total of 4, as you can see from the plot – this includes the highest peak)

Now, ok, there's 4 peaks, but maybe they are all very close to the highest peak and it doesn't matter, right? I mean, if the highest peak is, for instance, 90 and there’s 3 more peaks which only differ by maximum, say, 2 points from the highest peak, I would think that it’s not really a big issue.

That's were **maxFreqDiff** comes in

This tells you what's the maximum distance between the selected peak and the farthest frequency

that is, this:

Chart, histogram

Description automatically generated

In other words, here the algorithm is saying that the heart rate of this person is around 57. But there's 3 more peaks and the farthest is 18 points away...that's quite a higher HR. So, this doesn't look very trustworthy.

Finally, the **Flag system**

1. If there's more **than 1 peak beyond the threshold**, the video will be flagged with a **1**
2. If this criterion is met but, ADDITIONALLY, the furthest peak **is more than 10 points away**, this gets flagged with a 2

This should make it very easy to select these videos in R or maybe to go check their plots.