

## Document scope:

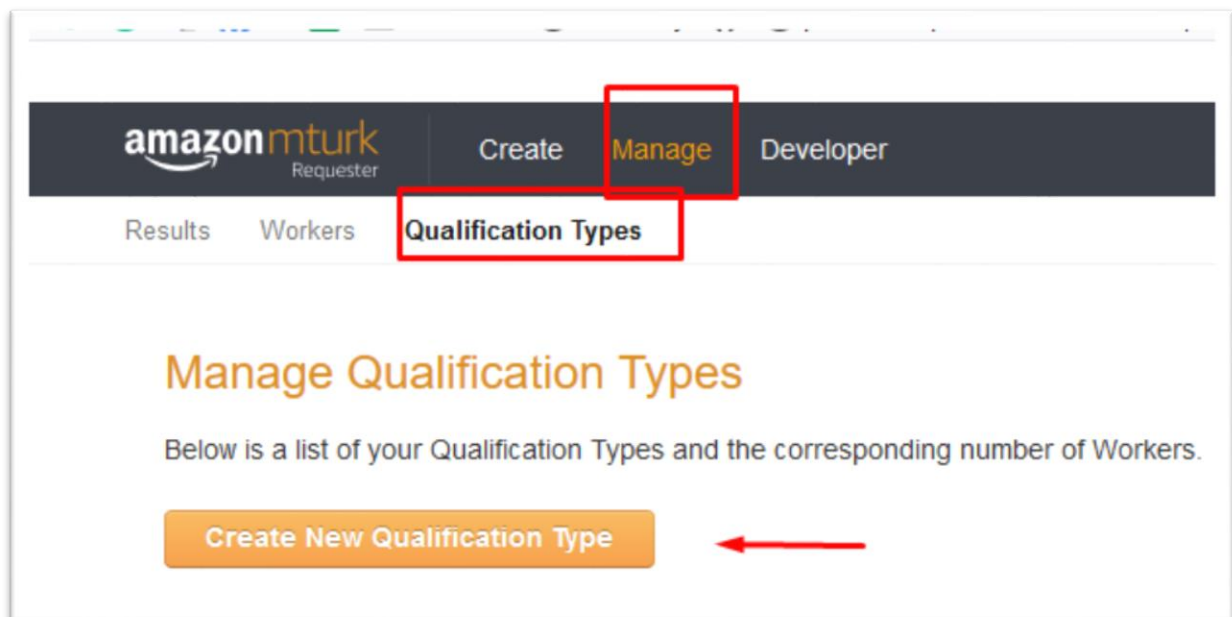
provide a brief set of instructions how to go from Batch\_3824355\_batch\_results\_output.csv to publishing the batch

### Step 1: Assign custom qualification to the accepted workers

**Step.1.1. Create custom qualification** (Note: one-time step, if qualification exist, you do not need to do this step).

From requester account:

- Go to “Manage” > “Qualification Types” > “Create New Qualification Type”



- Add a name: “ACR\_Listener” and description: “XYZ” (Note: workers see descriptions, so write something like “You are qualified to perform Listening Only Test -ACR”). And click OK.
- After a while (30sec), it will show up in the list
- You have created the qualification type.

### Step1.2. Assign the existing qualification type to accepted workers

From requester account:

- Open the “Batch\_3824355\_batch\_results\_output.csv” file and edit it to have following fields: (change WorkerId → Worker ID, Add a column “UPDATE-ACR\_Listener”, and add value 100<sup>1</sup> for all workers in that column). Save the file.

---

<sup>1</sup> Any other number is also working, just use the same number in step3 - Worker requirements

	A	B	C
1	Worker ID	UPDATE-ACR_Listener	
2	AMFTG54XC694N	100	
3			
4			
5			
6			
7			
8			

- Go to “Manage” > “Workers” > “Upload CSV”

amazonmturk Requester

Create **Manage** Developer

Results **Workers** Qualification Types

### Manage Workers

The Workers who have completed work for you are listed below. Select a Worker ID to bonus, block, unblock, assign a Qualification, or revoke a Qualification. To block, unblock, or change Qualification settings for multiple Workers, select Download CSV. Select Customize View to change which Qualification Types are displayed in the table below.

Customize View Download CSV **Upload CSV**

Show my Workers by: Lifetime Last 30 days Last 7 days

← Previous 1 2 3 4 5 6 7 8 9 ... 153 154 Next →

- Use the “Browse” button and upload the new “Batch\_3824355\_batch\_results\_output.csv”
- Click on “Yes” to assign the qualification.

#### Manage Workers > Processing File

### Processing File

Please review the following information and confirm

- **Assign 2 Qualification Scores**
- Revoke 0 Qualifications
- Block 0 Workers
- Unblock 0 Workers

Would you like to continue?

Cancel

Yes

## Step2: Prepare prerequisites

In this step you create and/or upload clips and other assets required for creating the HIT. All assets should be uploaded in a server and made publicly available. It is recommended to use any CDN service as well to make sure data are reachable for your participants.

### Step 2.1. Create trapping stimuli

Use script “create\_trapping\_stimuli” from hitapp\_p808/Scripts:

- **Edit** “hitapp\_p808/Scripts/cfgs\_and\_inputs/trapping.cfg” if needed
  - **“input\_directory= trapping”** it means all information will be find in a “trapping” directory. The “trapping” directory should contain three sub-directories: 1. messages, 2. source, 3. output. The directory path is relative to the script’s path. As a result, the script will look for following structure:

```
.
+-- create_trapping_stimuli.py
+-- trapping
|   +--messages
|       +--ACR_Bad_short.wav
|       +--ACR_Poor_short.wav
|       +--ACR_Fair_short.wav
|       +--ACR_Good_short.wav
|       +--ACR_Excellent_short.wav
|   +--source
|       +-- clip1.wav
|       +-- clip2.wav
|       +-- ....
|   +--output
```

- Add some clips from the dataset understudy in to ‘trapping/source’ folder. Make sure they include samples from every speaker, and different quality levels.
- Make ‘trapping/output’ directory empty: the trapping dataset will be created here
- **Run the script:**
  - **First check if requirements are installed:**  
pip install -r create\_trapping\_stimuli\_requirements.txt
  - **Run the script**  
python create\_trapping\_stimuli.py --cfg  
cfgs\_and\_inputs/trapping.cfg
- Check the trapping/output directory: 5 clips per each source clip should be created. In addition, you can find list of clips and their correct answers in “output\_report.csv”.
- 

### Step 2.2. Upload all resources

Create a csv file “row\_input.csv” and fill it: use a template given in ‘hitapp\_p808\P808Template\test input\row\_input\_template.csv’

- trapping dataset
  - upload the trapping dataset into your server.
  - Insert their URLs into the column “trapping\_clips” of “row\_input.csv”
  - Insert their corresponding correct answer into column “trapping\_ans” of “row\_input.csv” (Note: you can find the correct answers in output\_report.csv which was generated by “create\_trapping\_stimuli” script)

- Rating clips:
  - Upload your complete dataset into your server
  - Insert their URLs into column “rating\_clips” of “row\_input.csv”
- Gold standard clips:
  - Select and upload your gold-standard clips into your server
  - Insert their URLs into column “gold\_clips” of “row\_input.csv”
- Math questions:
  - Upload all clips in ‘\P808Template\assets\clips\math’ into your server
  - Insert their URLs into column “math” of “row\_input.csv”
- Environment Test (setup section /pair comparison)
  - Upload all clips in ‘\P808Template\assets\clips\sample\_jnd’ into your server
  - Insert URLs of files ‘50\*.wav’ in ‘pair\_a’ and ‘42\*.wav’ into column “pair\_b” of “row\_input.csv”
  - Note: each row should match i.e. belonging to same speaker e.g. ‘42S\_female1.wav’ and ‘50S\_female1.wav’
- Training
  - Upload your training clips into your server
  - Insert their URLs into the ACR.html file (var config[‘trainingUrls’]).
  - In case you want to have a trapping question in the training:
    - Insert its URL also into the config[‘knownQuestionInTrainingUrl’]
    - Insert its correct answer into the config[‘knownQuestionInTrainingAns’]
- Other resources
  - Following resources should be also uploaded into a server and their URL should be changed in the ACR.html
  - Volume setting: “hitapp\_p808\P808Template\assets\clips\signal\_level.wav”
  - Image in Instruction: “hitapp\_p808\P808Template\assets\img\process\_2.png”
  - Image in setup: “hitapp\_p808\P808Template\assets\img\attention.pn”

### Step 3: create input.csv

- Use script “create\_input\_acr” from hitapp\_p808.

The script needs two input files:

- A configuration file. Example is given in “Scripts\cfgs\_and\_inputs\create\_input.cfg”
- row\_input.csv which was create in Step2.

The row\_input file should contains following columns:

- 'rating\_clips': urls of all clips which needs to be rated
- 'math': url of various math questions to proof usage of two-eared headphones.
- 'pair\_a','pair\_b': pairs will appear in the setup section, to check the environment of user.
- 'trapping\_clips','trapping\_ans': url to all trapping questions, and a number which shows the correct answer.
- 'gold\_clips': (optional) list of gild clips
- Run the script:
  - First check if requirements are installed:  

```
pip install -r create_input_acr_requirements.txt
```

- **Run the script**

```
python create_input_acr.py --cfg cfgs_and_inputs/create_input.cfg
--row_input cfgs_and_inputs/ row_input_librivox.csv
```

The script generates an input csv file: xxx\_publish\_batch.csv

#### Step4: Create the ACR Project

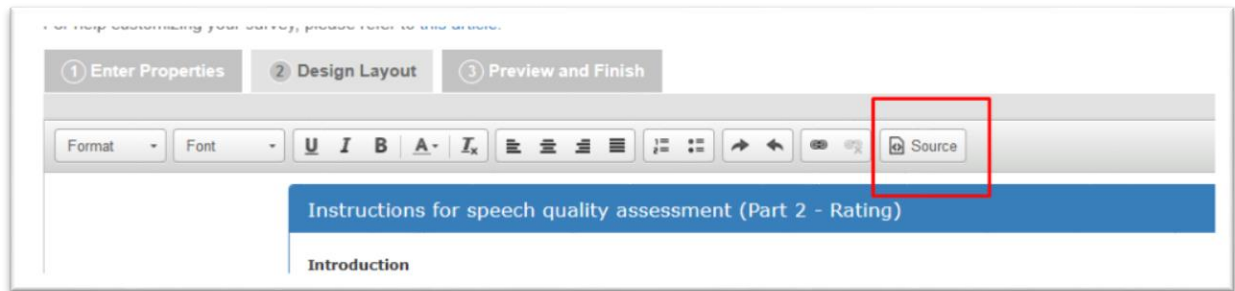
- Go to “Create” > “New Project” > “Survey Link” > “Create project”

The screenshot shows the Amazon Mechanical Turk 'Create Project' interface. In the left sidebar, 'Survey Link' is highlighted. The main content area displays 'Survey Link Instructions (Click to expand)' with a 'Survey link' field containing 'http://example.com/survey345.html' and a 'Provide the survey code here' field with 'e.g. 123456'. A blue button indicates 'You must ACCEPT the HIT before you can submit the results.' At the bottom right, there is an orange 'Create Project' button.

- Fill information in “1 – Enter Properties”, important ones:
  - “Setting up your survey”
    - **Payment**
    - “Number of respondents”: **9**
    - “Time allotted per Worker”: **1 Hours**
  - “Worker requirements”
    - **Use following qualifications:**

The screenshot shows the 'Worker requirements' section. It starts with the question 'Require that Workers be Masters to do your tasks (Who are Mechanical Turk Masters?)' with radio buttons for 'Yes' and 'No' (selected). Below, it says 'Specify any additional qualifications Workers must meet to work on your tasks:'. There are two qualification rows: 'HIT Approval Rate (%)' with a dropdown set to 'greater than' and a value of '98', and 'ACR\_Listener' with a dropdown set to 'equal to' and a value of '100'. Each row has a 'Remove' button. At the bottom, there is a button to 'Add another criterion' with a note '(up to 3 more)' and a link to 'Pricing Details'.

- Save and Go to “Design Layout”
- Click on “Source”



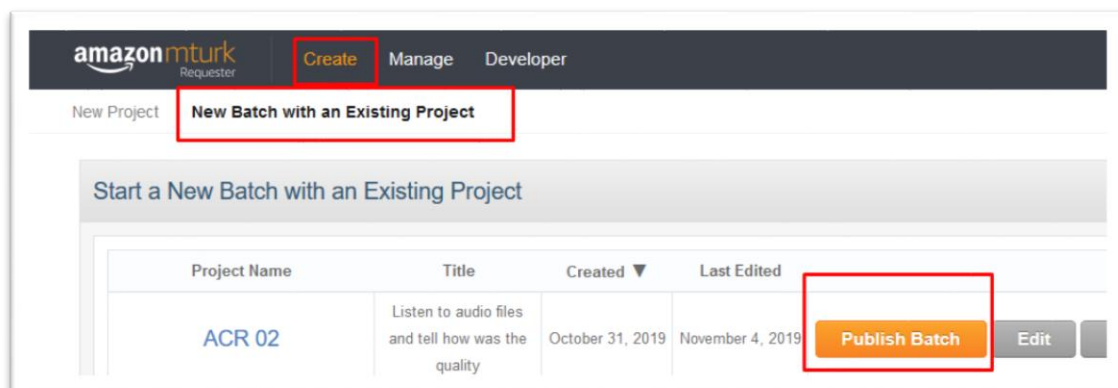
- Copy-paste the ACR.html here
  - NOTE: you should consider to change some information in the “Introduction” section: like **“Payment”, or the image on top (up to 60x Ratings) ...**
  - Note: you may consider to change the config object as well- variables should match with input.csv



- Click on “source”
- Click on “Save” and “Preview”
- Click on “Finish”

#### Step 5: Publish the batch

- Go to “Create”> “New Batch with an Existing Project”



- Find your project, and click on “Publish Batch”
- Upload your input.csv file created in last section (xxx\_publish\_batch.csv)

Publish Batch

Choose a .csv file with the variables you specified in your project. ([learn more](#))

Choose File

row\_input\_libri...\_10selected.csv

Upload

File validation completed successfully

Validated

File Name: row\_input\_librivox\_publish\_batch\_10selected.csv

File Validated: Yes

File Size: 17.55 KB

Line Count: 11

Don't have a data file? [Download a sample .csv file.](#)

- If everything is green, go ahead with “upload” otherwise there is variable missing in the input.csv.
- Check the HITs, and if everything is ok click on Next

Listen to audio files and tell how was the quality

Requester: babak

Reward: \$0.05 per task

Tasks available: 10

Duration: 1 Hours

Qualifications Required: ACR\_Listener equal to 100

Instructions for speech quality assessment (Part 2 - Rating)

Introduction

Qualification HIT (3 minutes)

Up to 60x Rating HITs (each 5 minutes)

Welcome and congratulation! You have been selected to participate in our speech quality assessment experiment! This HIT has two + one (just every one hour) sections::

- **Setup:** Configure your system and validate it by answering to 6 questions
- **Training (every one hour):** questions, same as "Ratings" but appears just once a day
- **Ratings:** Listen to audio files and give **your opinion about the quality of the speech** you hear.

You should follow the below mentioned rules, otherwise your answers will be invalid.

**Rules:**

- Use a headset, not the loudspeaker: otherwise your response will be rejected
- Perform the task in a quite environment
- Do not change the volume after modifying it in the Setup section.

**Payment:**

Next HIT

Cancel

Next

- Check the calculation and “Publish”
- You may send emails to workers to inform them about availability of this project (see next page)

## **Sending Emails to Workers:**

- Edit the configuration file 'Scripts\mturk.cfg'
  - Add 'aws\_access\_key\_id', and 'aws\_secret\_access\_key' of your requester account (if you do not have them, follow Step1 and Step2 here: <https://requester.mturk.com/developer>)
  - Make sure the 'endpoint\_url' of production is active and the one for sandbox is commented
  - Add your "requester\_client\_id" (login as a worker to see it)
  - Edit the "subject", and "message" (note: the url in the message will refer to a page containing all HITs created by your requester\_client account)
  - Add a comma separated list of "worker\_ids"
- **Run the script:**
  - **First check if requirements are installed:**  

```
pip install -r mturk_utils_requirements.txt
```
  - **Run the script**  

```
python mturk_utils.py --cfg mturk.cfg send_emails
```