**Meeting 12/11/20 Notes**

* Progress halted as can’t get permission/funding to use Mechanical Turk – working on it
* Plan B – can eventually turn back into plan A after finally get annotations
* Split dataset with demographics of annotators into test and train and see if can predict demographic of annotator based on toxicity score – e.g. for comment + high toxicity score is it more likely to predict female than same comment + low toxicity score (can do same for other demographics to find bias)
* Shed more light on specific issues – sexism, racism
* Next target: statistical analysis – correlation heat map, coefficients between gender across toxicity scores, toxicity scores categorical – divide scale into bins and get distribution according to gender, plot demographic information, heat maps