

Research Question:

1. How does annotation behavior change throughout annotation task?

2. Does order affect annotation likelihood?

→ 3000 tweets annotated for **Hate Speech (HS)** and **Offensive Language (OL)**

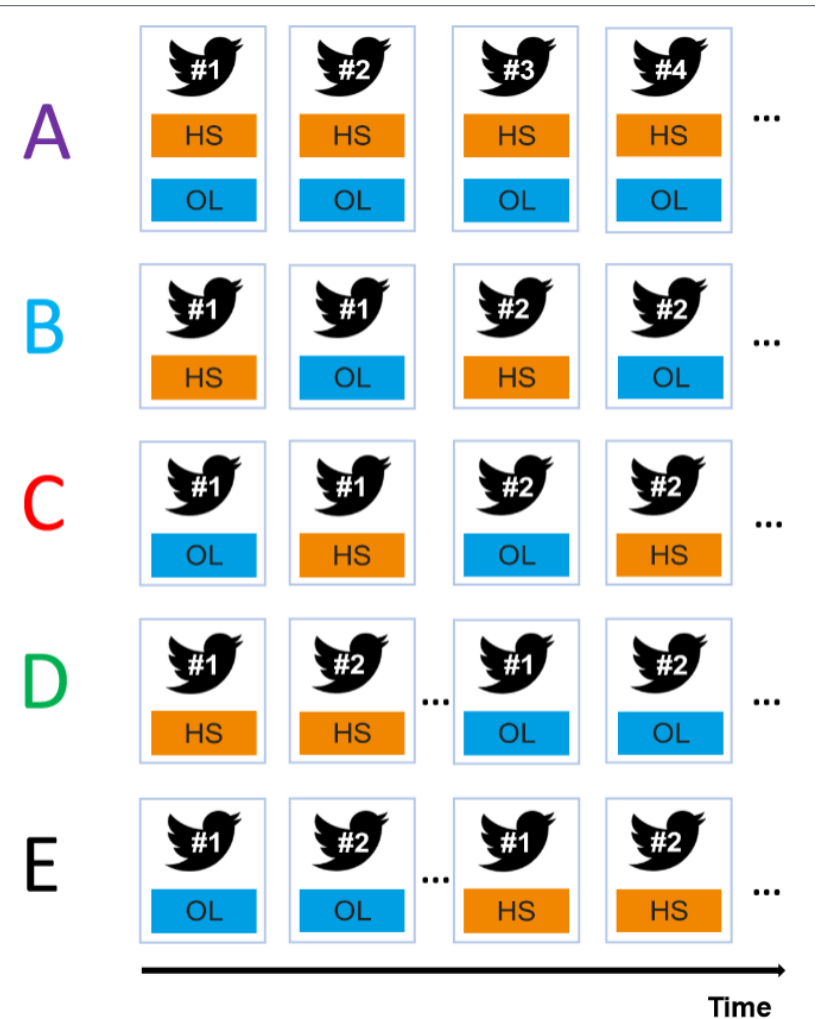
→ 3 annotators in each of the five exp. conditions per tweet

→ Tweets grouped in batches of 50.

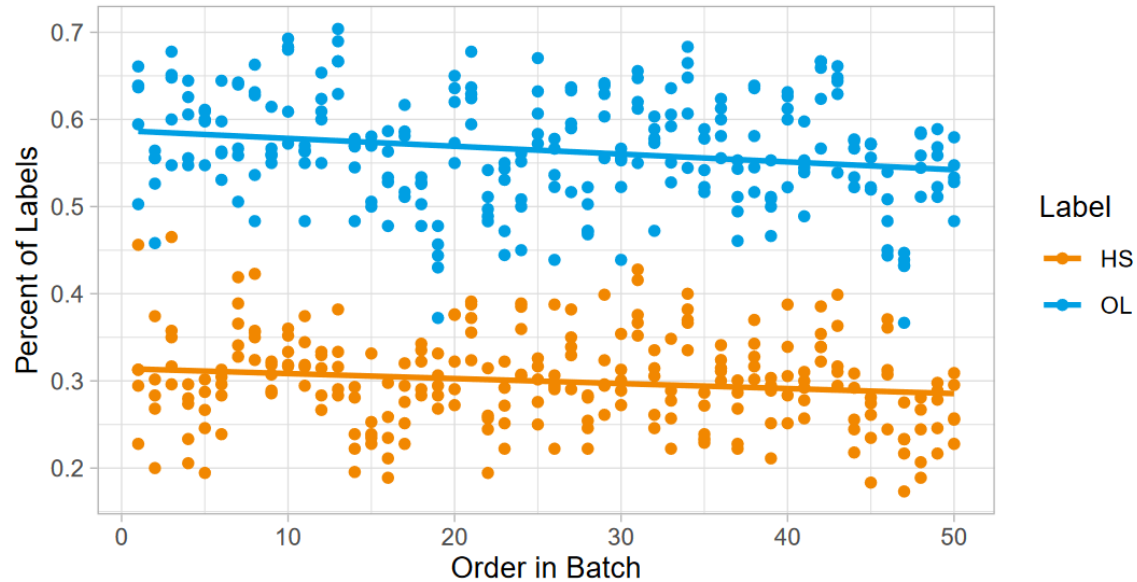
Each individual annotated one batch

→ Order was constant within batches

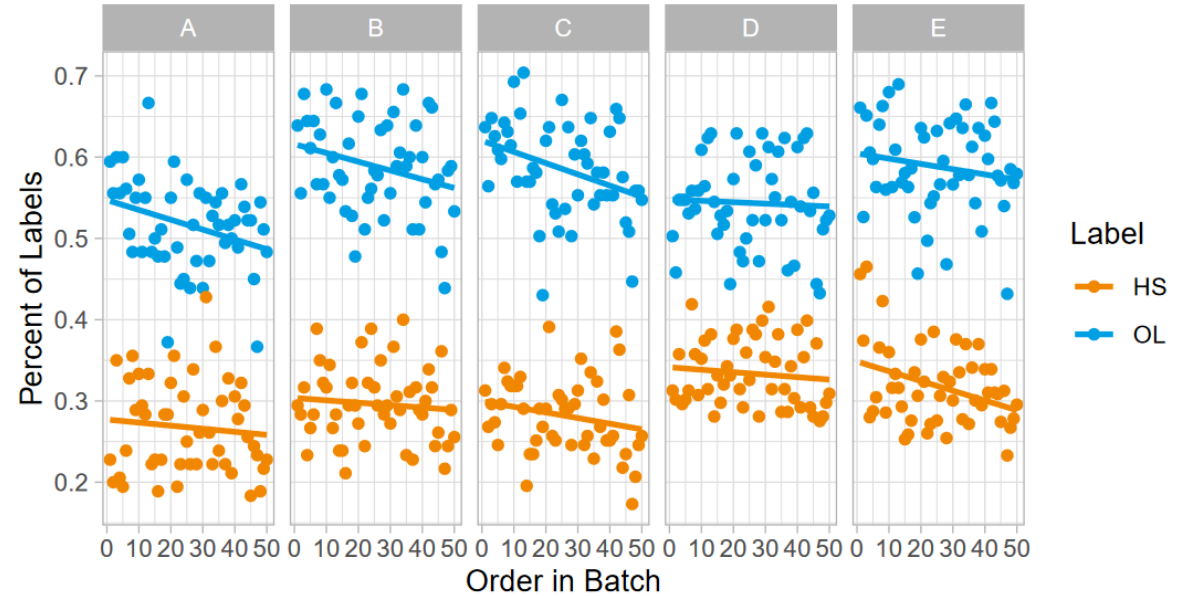
→ Evaluation: Linear probability regression models



Regressions for full data



Regressions by exp. Condition



→ The likelihood for HS/OL annotation never increases throughout task; often negative effect of order

→ Slopes steeper for OL

→ Non-random annotation instance ordering threatens data quality! (e.g. Active Learning)

→ Let's discuss our results in the poster session!