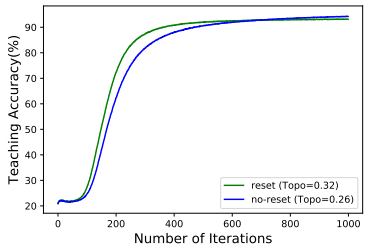

Figure 1: Ease-of-teaching.

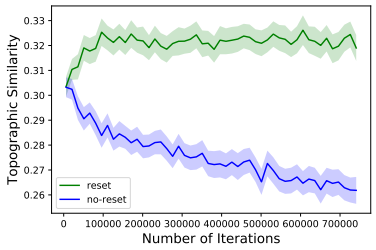

Figure 2: Topographic similarity.

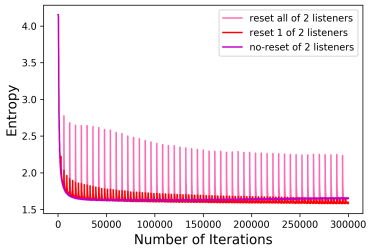

Figure 3: Speaker's entropy ($N = 2$).

First, we want to thank all the reviewers for the helpful comments and suggestions which will be reflected in future iterations.

R2: We have since ran further experiments with 3 attributes and 8 values per attribute, resulting in $8^3 = 512$ objects and a message length of 3. Sensibly, we increased the iterations between resets to give more time to learn in this more challenging communication task (15k instead of 6k). The same trends continue: see Figure 1 and Figure 2.

The primary contribution of the population experiments is to explore whether the observed improvements come from the sequence of agents acting as a population. We found instead it is rather the abrupt change in listener that matters for the effect. The pattern of the entropy term for non-staggered resets does match the single-agent reset case (see Figure 3 for the $N = 2$ case). We will make both of these observations more clear and include such graphs to strengthen this hypothesis as suggested.

Related works in iterated learning should be discussed more thoroughly. We will change the wording and include a discussion of the papers you noted.

An initial investigation into the degree of language change suggests that the amount of change in the language during training is similar across all regimes (with and without reset), with as much as half the language changing in 6k iterations.

We had a transcription error in Table 1 in the paper so the whole third column of the table should read "be ee ce ge". The language shown is to compare the most structured languages we could find after training 300k iterations, with ones from the no-reset regime.

Finally, we want to confirm that your other assumptions in "minor comments and clarifications" are correct and will endeavour to clarify these points in the text: all the listeners are reset to random parameters; we calculated the entropy term with a token-level approximation; for ease-of-teaching evaluation the teacher is frozen; in the population experiments ease-of-teaching is evaluated after 300k iterations; "reset 1" needs to be clarified and we will use your "staggered" nomenclature; the results of "simultaneous resetting" in $N = 1, 2, 10$ are all similar and statistically inseparable.

R3: We should be clear that the reviewer's third described contribution — the metric of topographic similarity — is not novel to this work. On the other hand, the metric of ease-of-teaching is a new metric of evaluation and we think it is a valuable contribution to the emergent communication field.

Thanks for pointing out several avenues this work opens. We also think explicitly optimizing ease-of-teaching is a worthwhile direction of future work.

R4: We will add a brief discussion of the papers you mentioned. Note that communication actions (i.e., speaking and listening) are the only actions in referential games. In such cases, successful task completion ensures "useful communication is actually happening".