Introduction

The raw data-set was split into several files, division being de facto based on order of acquisition. The experiment was run three times in total with one year gaps. No participant appeared more than once.

For our analysis however, we used only the most recent instance of the experiment (2018). In order to test whether there is physiological coordination, we compare heart rate and respiration data from the each pair of interlocutors via CRQA-analyses to acquire measures of coordination like the recurrence rate (RR) and the determinant.

These measures were also obtained for shuffled pairs and surrogate pairs. The shuffled pairs are created by scrambling the time series while the surrogate pairs are the complete set of possible combinations of individuals from the real pairs excluding the real pairs.

Because we expect the effects of the three conditions on coordination to differ between the different types of pairs, we hypothesise an interaction effect between condition and pair type. Specifically, we expect the shuffled pairs to have low and similar recurrence measures in all conditions. Recurrence for real pairs is expected to follow behavioral coordination and thus be highest in the synchronous task and lowest in the conversation task. Surrogate pairs are expected to follow the same trajectory across the three conditions as the real pairs, but have lower coordination values in the conversation condition because the conversation task is idiosyncratic to each particular real pair.

As we do not specify random structures in our model, we end up with the following general model structure:

```
lm(recurrence measure ~ condition * pair type, data).
```

Task 1, Preprocessing the data

We started by combining contents of provided raw files into a single data frame. Next, records were properly indexed and variables containing ECG measurements were removed. Furthermore, we downsampled our records, as the original number of rows seemed excessive for the purposes of this analysis.

Downsampling was achieved by dividing our observations into groups containing one hundred observations each. Next, means of available variables within these groups were calculated and resulting values were further used as a substitute for original data. This has lead to a reduction of our sample to 51,712 records, massively shrinking computational demands of later analysis.

Downsampled data was checked for outliers. Observations within the distance of 2.5 standard deviations from mean of their respective time series were left untouched. Observations outside these boundaries were changed to the mean of their respective time series. Notice that outlier removal was done across condition but within metric (HR, RESP).

Figure 19: Before (left) and after (right) outlier removal
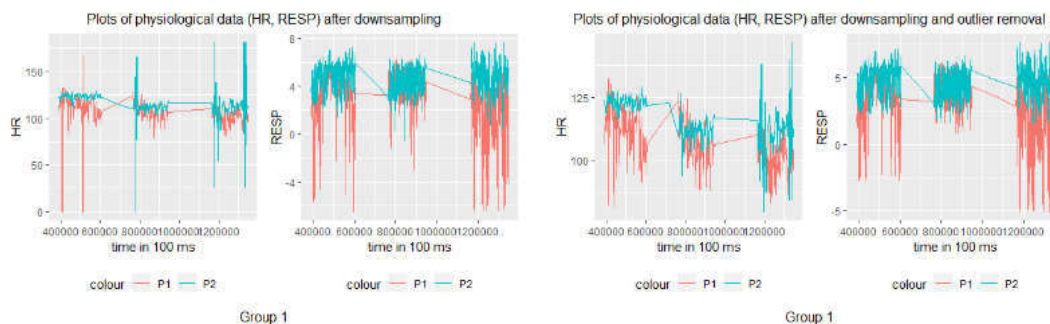
We made recurrence plots for each condition and for each physiological metric. As can be seen in the recurrence plots for respiration ('Figure 21'), there are few to no diagonal lines in any of the conditions indicating that the length of coordination is low in all tasks.

We do, however, see some vertical structure especially in the synchronous task. Overall, these plots are very hard to interpret as they do not display any clear structure.

The recurrence plots for heart rate ('Figure 22') reveal basically no recurrence between the interlocutors which might have to do with the parameters outputted from the optimzeParam() are not optimal.
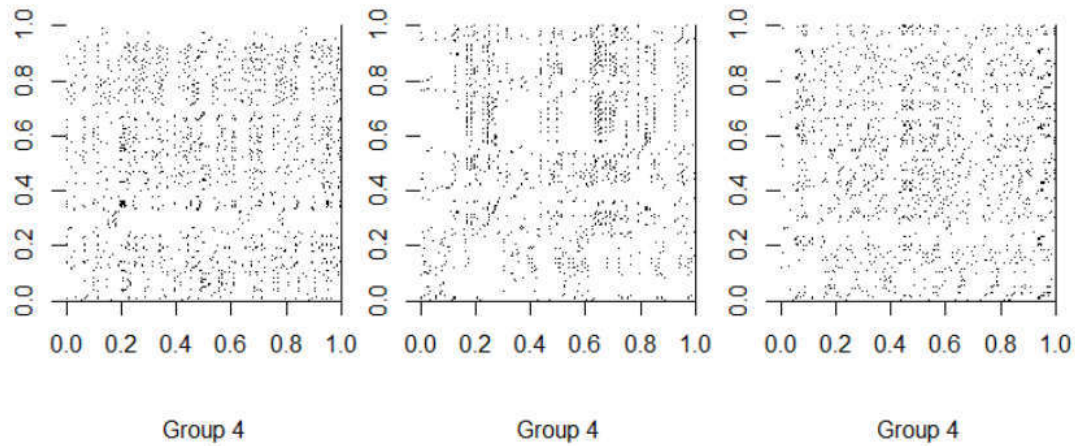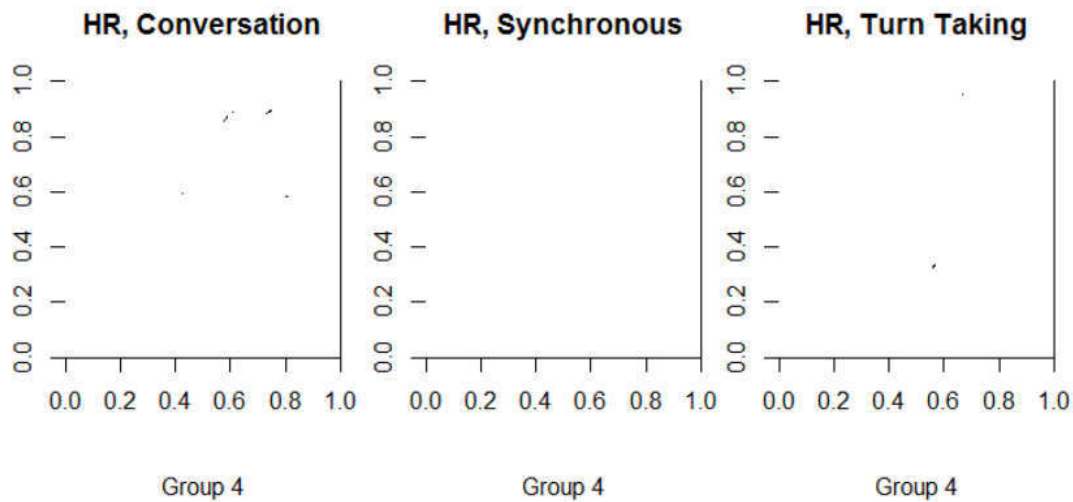
Figure 21: Recurrence plots for respiration

Figure 22: Recurrence plots for heart rate

Task 2 & 3, Models and Results

We tested whether there was interpersonal alignment & whether conditions differed in the same models.

The general structure of the models `lm(recurrence measure ~ condition * pair type, data)` was applied to recurrence measures RR, determinant, & L. In order to make separate models for the two physiological

Figure 24: Observed data for respiration versus model predictions for respiration, including shuffled pairs
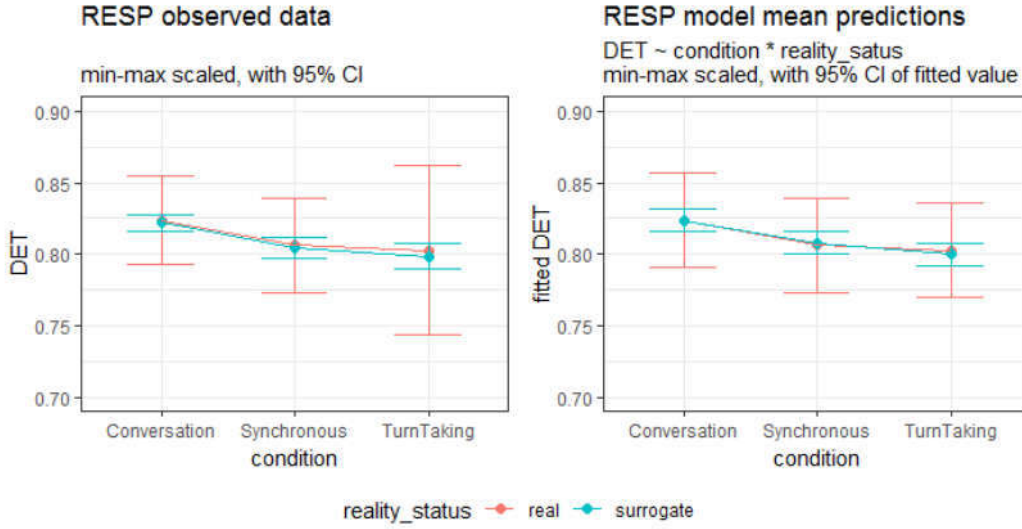
Task 4, Effect of Respiration on Heart Rate Coordination

There seems to be two ways of testing the effect of respiration coordination on heart rate coordination:

- First approach: build a linear (mixed effects) model in which the predictor is a measure of respiration coordination (RR, DET, etc.) and the outcome variable is the corresponding measure of heart rate coordination. One could potentially add something as random effects (maybe condition).

  - M1 ← lmer(hr1S.RR ~ resp1S.RR, data)
  - M2 ← lmer(hr1S.DET ~ resp1S.DET, data)

- Second approach: Use a machine learning approach in which we try to predict a measure of heart rate coordination (RR, DET, etc.) based on data of respiration coordination. Then one could plot the error of the estimate. The best approach would probably be to cross-validate. Alternatively, one could do the following which have been sketched in the R-document, but which does not work well as of now.

1. Train / Test indexing. We have used a subset of the data without shuffled because these contain almost exclusively zeroes.
   2. Test / Train set based on indexing. This should take into account that we probably want an equal percentage of different conditions and pair type in the training & test dataset (which is not done in the outline code provided).
   3. Building a model which predicts a feature of HR (e.g. DET) based on all features of resp (maybe only significant predictors) & perhaps with random intercepts for condition & reality status. The model should use the training data. See an example of such a model in the code. M1 ← `lmer(hr1S.DET ~ resp1S.AllFeatures + (1|condition) + (1|reality_status), train, REML = FALSE)`
   4. Use the model to predict values of hr1S in the test dataset
- `dataset$predictions ← predict(M1, test, allow.new.levels = T)`
   5. Compare the predictions to the actual values of the parameter. We have done a plot in the code which shows that the predictions (as of now) are useless.