Lab Assignment 1: Processing Models

Luka van der Plas (4119142) & Rianne Lam (6888216)

11 november 2019

Section I: Human data

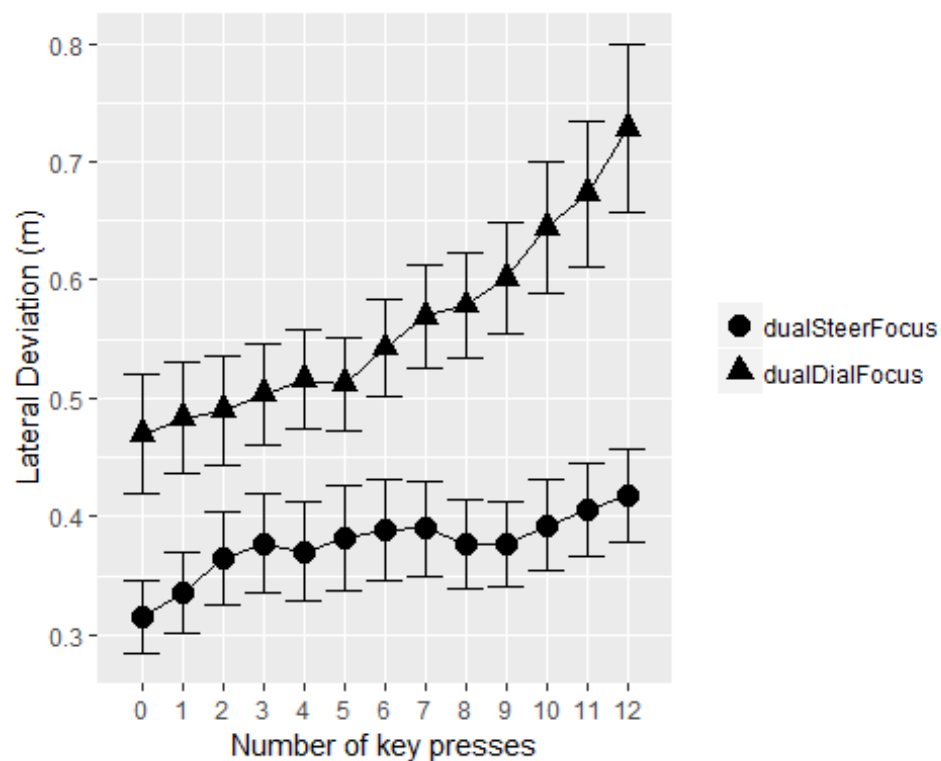### Question 1A

```
Condition       M       SD       SE
Steering focus 5068.79 1317.82 380.42
Dialing focus   3432.71  896.05 258.67
```

Question 1B

```
Condition       M     SD     SE
Steering focus 0.376 0.319 0.033
Dialing focus   0.562 0.376 0.037
```

Question 1C

Figure 1: Lane deviation over time.

Question 1D

For both conditions, it seems that participants did not interleave dialing for driving between the 5th and 6th key presses. At the 6th key press, the absolute lane deviation did not decrease compared to the 5th key press, while this would have been expected if participants were to focus on the road at this point.

However, in the dialing focus condition we can see that from the 6th key press onwards participants tend to focus less on the road and thus show greater absolute lane deviation. This gives a kind of breakpoint in the plot. But this result cannot be explained by the hypothesis that particpants would interleave dialing for driving at this point in time.

Section II: Fitting lateral drift of the car

Question 2A

Figure 2: Changes in the lateral position of the car over time (between 15 and 18 seconds).
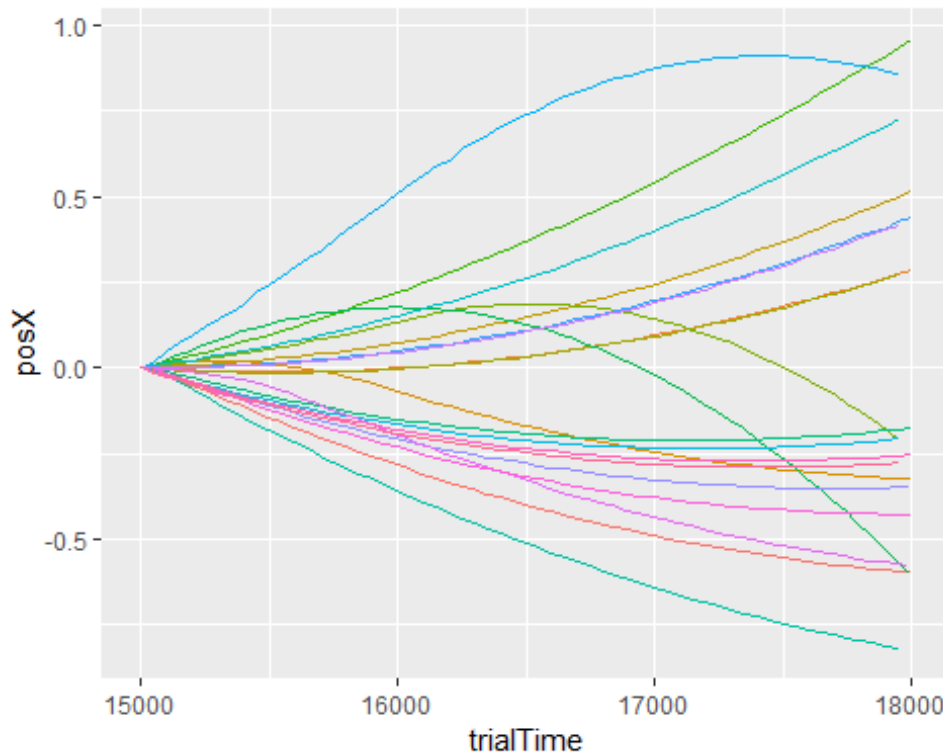
Question 2B

Figure 3: Changes in the lateral position of the car over time (between 15 and 18 seconds) for 20 simulated trials.

Question 2C

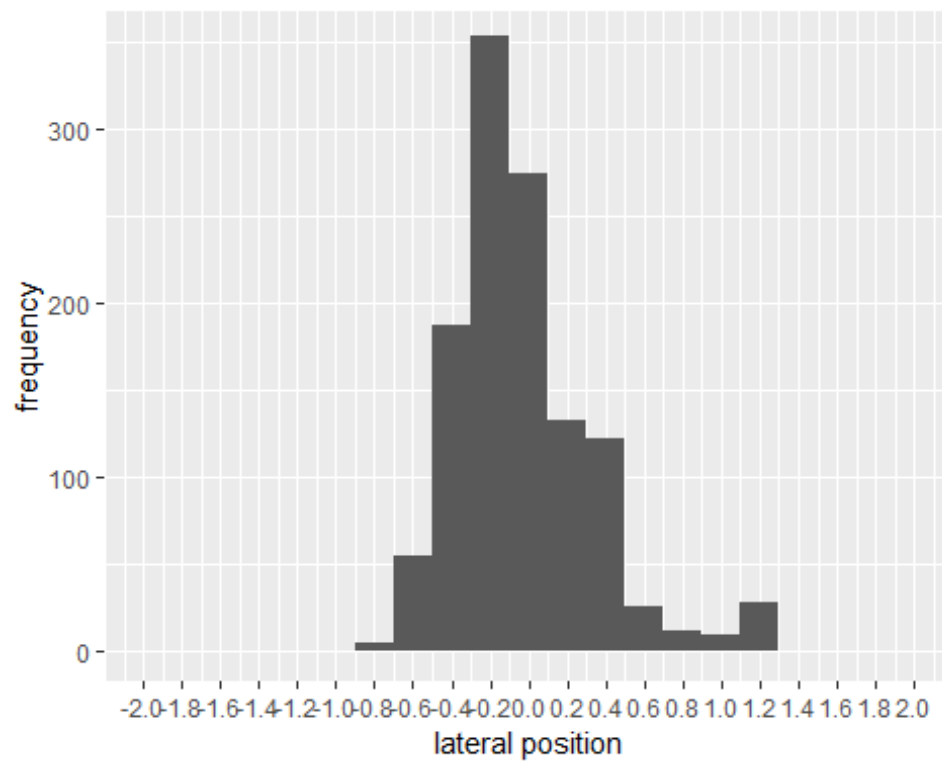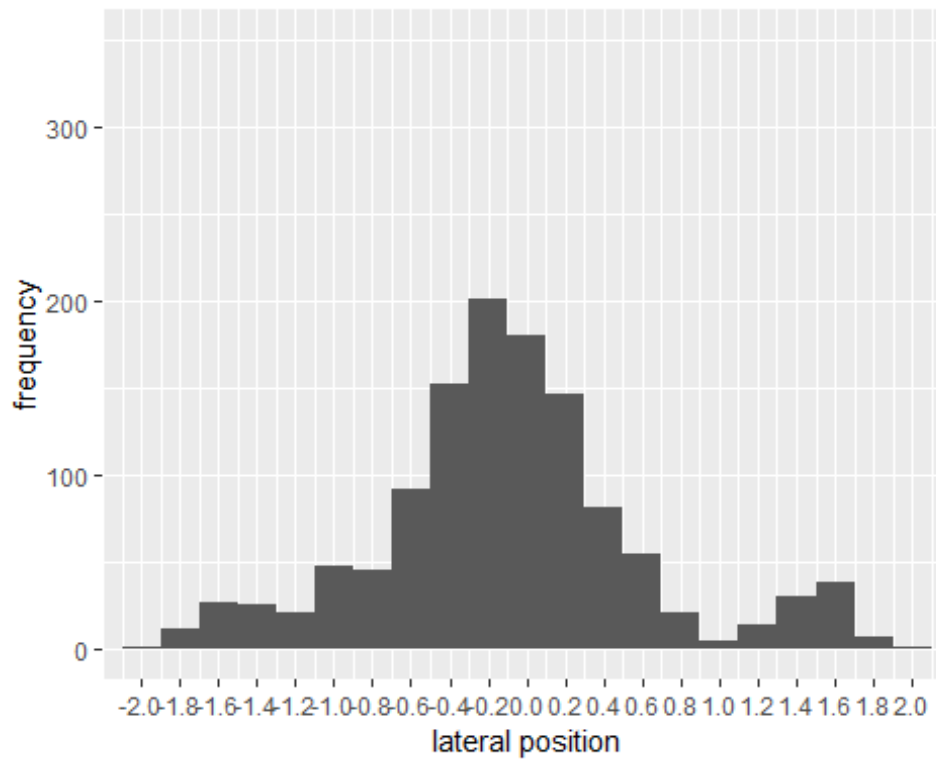Figure 4: The distribution of car positions as measured in the human trial.

Figure 5: The distribution of car positions based on the simulated data.

Question 2D

```
Source      SD
Measured    0.3596126
Simulated   0.6934516
```

Question 2E

To find the optimal SD value, we searched through SD values from 0.01 to 0.13, with a step of 0.01. For each SD value, we calculated the standard deviation of the lateral position over 50 trials, and then calculated the absolute difference with the target SD, i.e. the one measured on the human trials. Below is a plot (Figure 6) with these results.
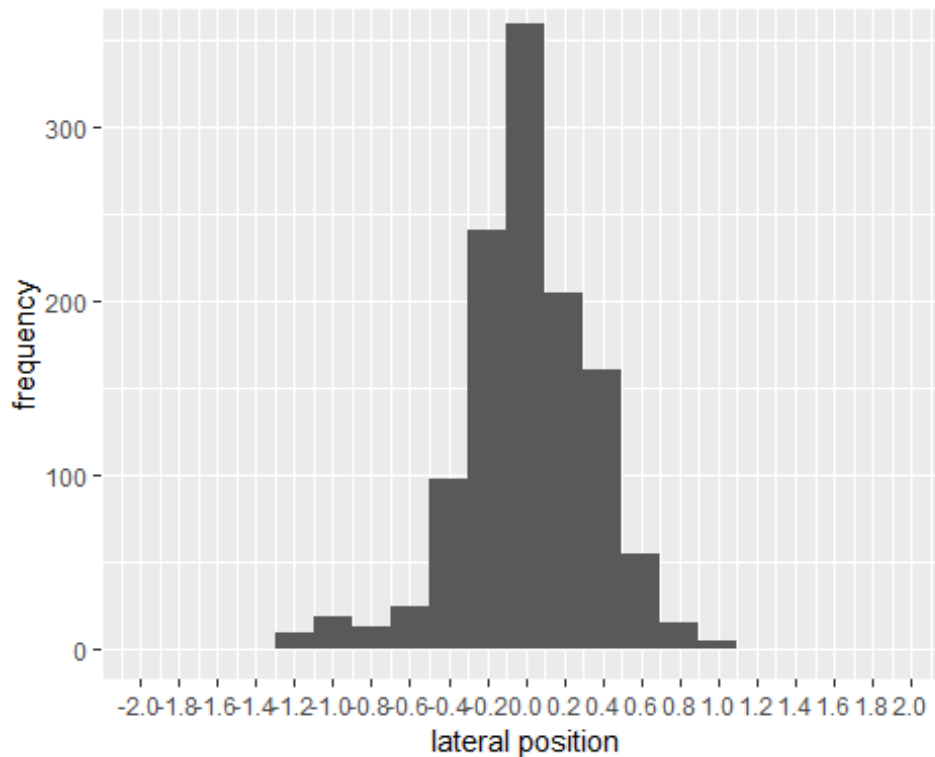
Based on these results, we chose to use an SD for the gaussian distribution of 0.07. This gives an SD of lateral position of 0.344.

Figure 7: Changes in lane position over time for the individual simulated trials.

Figure 8: The distribution of car positions based on the simulated data using the new SD value.

Section III: Fitting Interkeypress intervals (IKIs)

We changed the gaussDeviateSD in the driving model to 0.07 for drifting, and the gaussDriveNoiseSD to 0.053, to preserve the ratio between the two parameters.

Question 3A

The average keypress time is 272.5 ms.

Question 3B

We picked a value of 275 ms, because that amount of rounding seemed acceptable.

Section IV: Expanding the strategy analysis

Question 4A
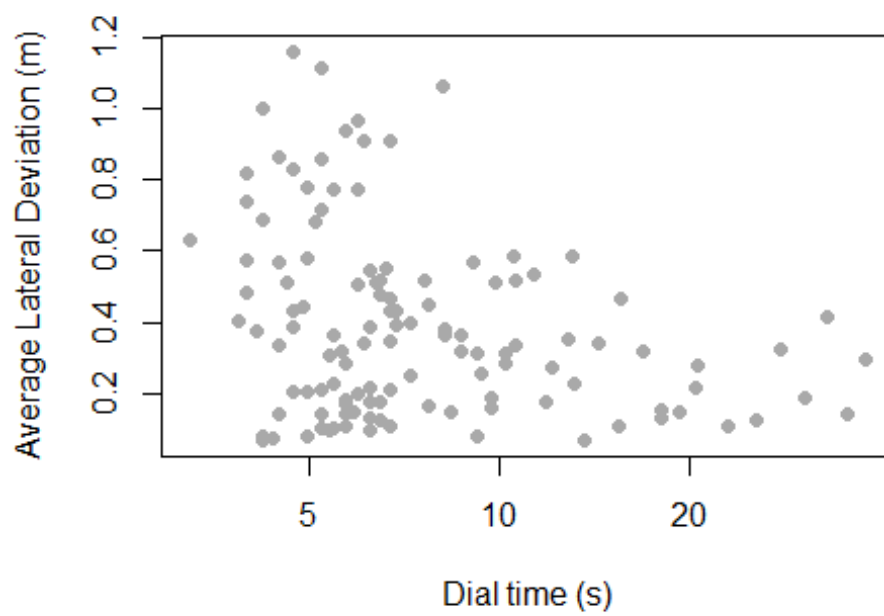
Figure 9a: 1 simulation per strategy.
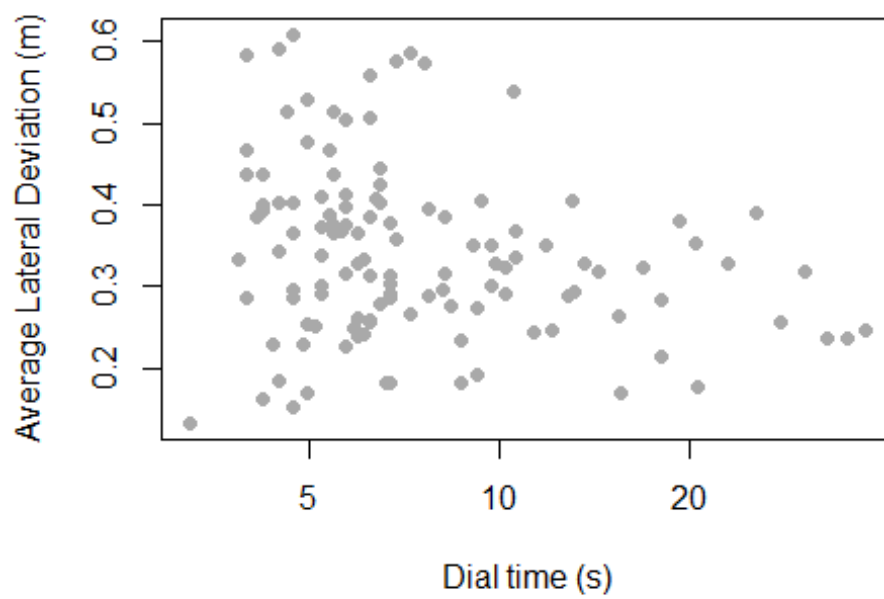
Figure 9b: 5 simulations per strategy.

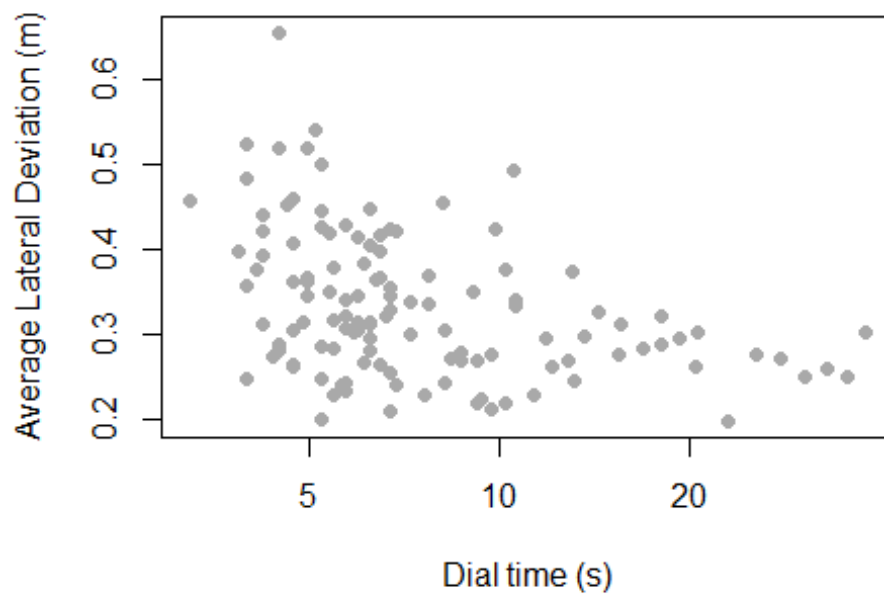Figure 9c: 10 simulations per strategy.
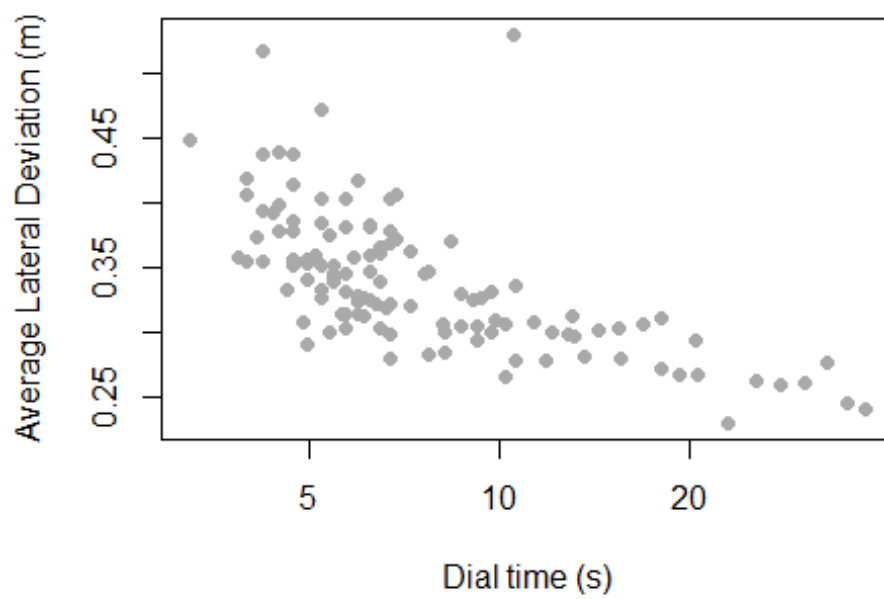
Figure 9d: 50 simulations per strategy.
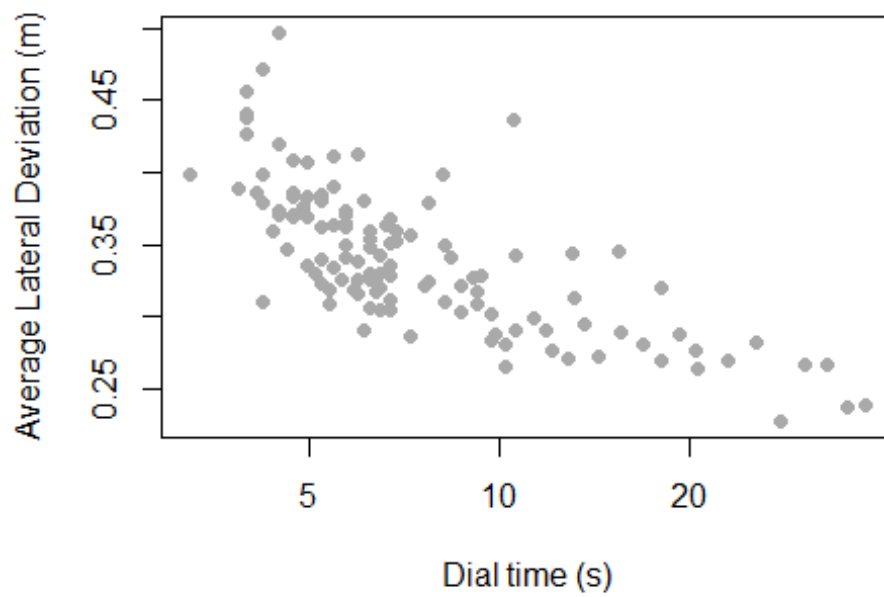

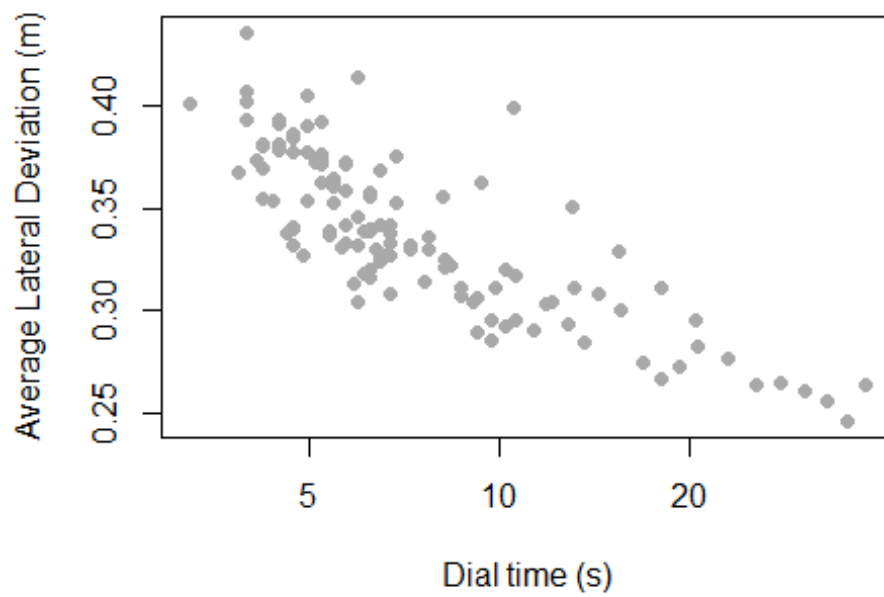Figure 9e: 100 simulations per strategy.

Figure 9f: 200 simulations per strategy.

Question 4B

The data show that results become less noisy with more simulations: the range of total deviation is smaller and data show a clear downward trend between deviation and dial time. We wanted to quantify this result: a clearer trend with fewer outliers will give a stronger correlation between dialing time and deviation. This correlation is plotted for all simulation counts below.
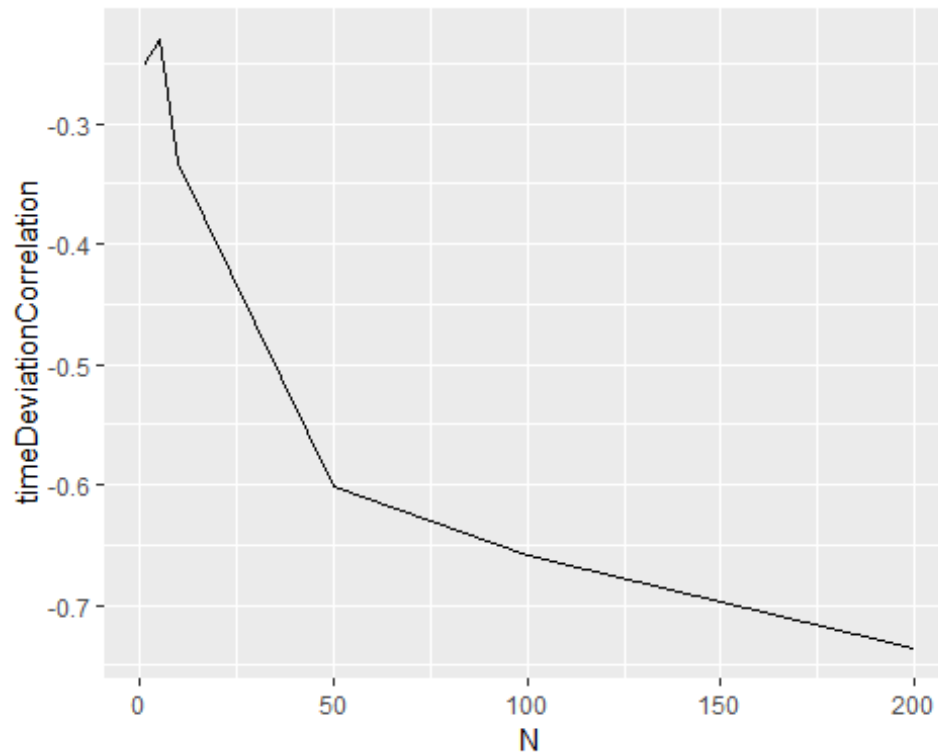

We estimate that the 50 simulations should work well in development: they do not take a very long time to run, but already show a clear trend in the data. The correlation plot shows that after 50 simulations, the amount of trend is not changing that much anymore. If time allows, it may be good to run 100 simulations instead of 50, since there is some increase in correlation (and therefore a decrease in the noise) between 50 and 100 simulations. However, there is almost no increase between 100 and 200.

Question 5A

To run simulations from this point on, we adjusted the steering time options to (2,4,6,8,10,12).

Figure 10: Simulation results of the revised model together with human data.
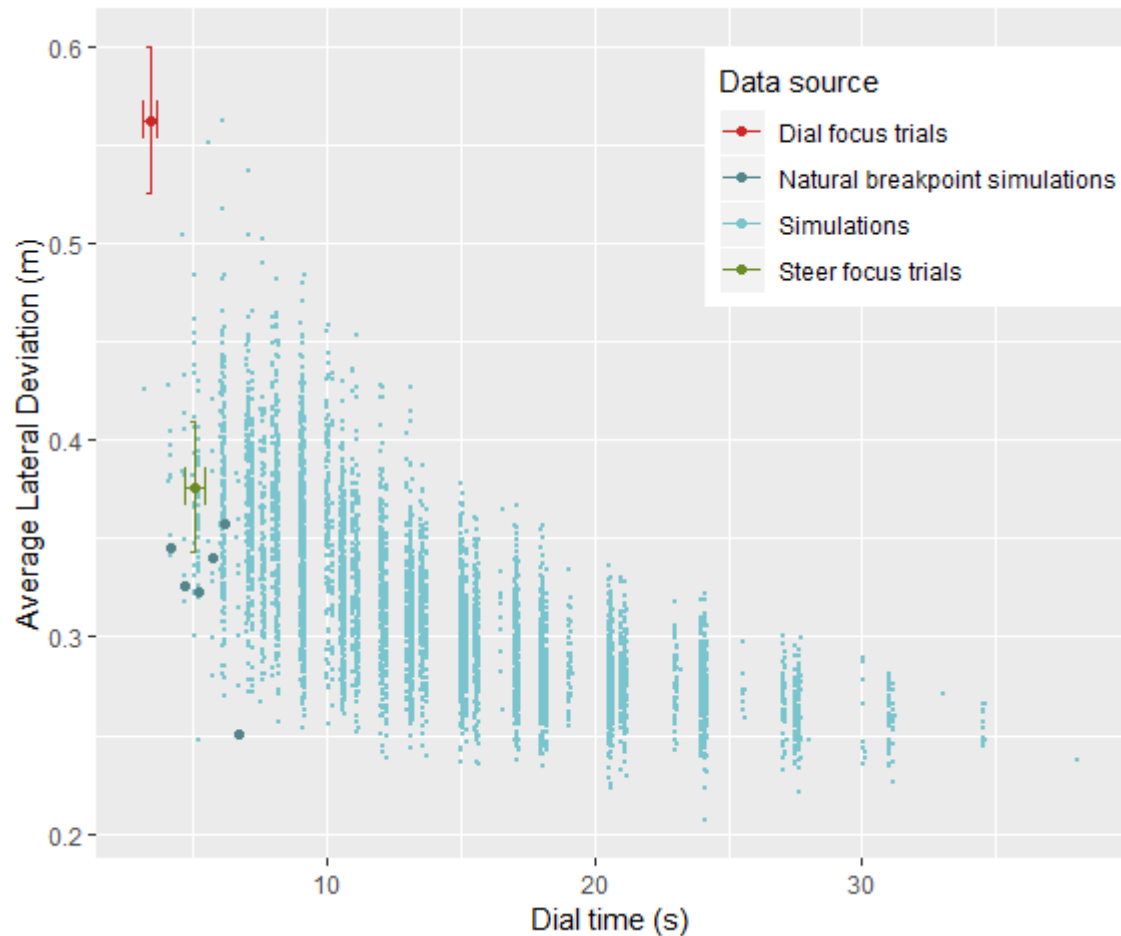
Question 5B

In the dial condition, the human performance falls outside of the range of the model. However, in the steering condition, it can be said that the human performance is on the outside edge of the performance space. Specifically, the human data fall in the left/lower boundary of the performance space for the model, matching with the models that take relatively little time given their deviation, and make little deviation given their trial time. The strategies in these simulations had the most efficient performance, so if we assume that the model is valid, it can be concluded that humans must also use their time efficiently.

Question 5C

We selected the simulations for which the data fall within the margin of error of the steering condition for the human trials. This gave 30 strategies. Below, the breakpoints for each of these strategies is displayed (Figure 10).
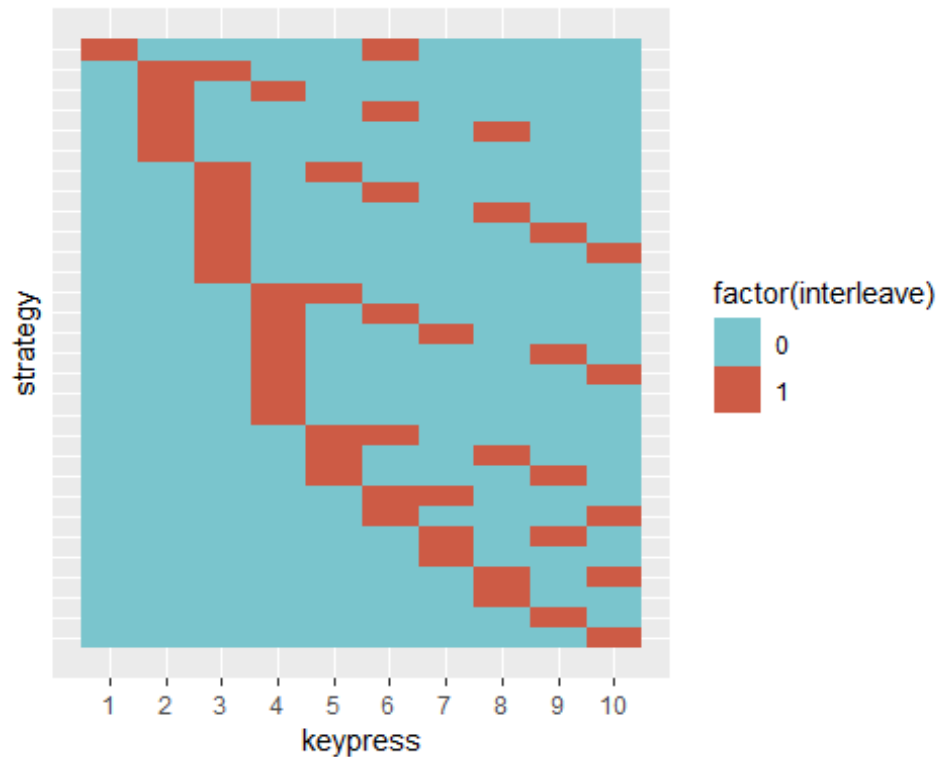
A few things to note:

* The first breakpoint can be anywhere from the first to the tenth digit. However, there are slightly simulations where the first breakpoint was in digits 2-4.

* All of the strategies have only one or two breakpoints.

* There does not seem to be a bias towards any particular position for breakpoints (like the 5th digit).

* Strategies that interleave solely between the 5th and 6th digit are not in this selection.

We can definitely not conclude that humans must favour the "natural breakpoint" between the 5th and 6th digit, since the results for those simulations do not match with human data. However, the data do suggest that humans may take only one or two breaks, where the first break is often after keypress 2, 3, or 4.

Section V: Relationship between model flexibility and model fit

Question 6

Because all human-like results came from simulations with only one or two breakpoints, we hypothesised that the number of simulations can be reduced by leaving out the strategies with a very large number of breakpoints. If humans make very few breakpoints, there is no reason to run simulations for a lot of potential strategies that do not match human performance.

To test this hypothesis, we separated the data from question 5 on whether there were at most 5 breaks. These results are plotted below (Figure 12).

Figure 12: Model prediction separated by whether there are at most 5 breaks.
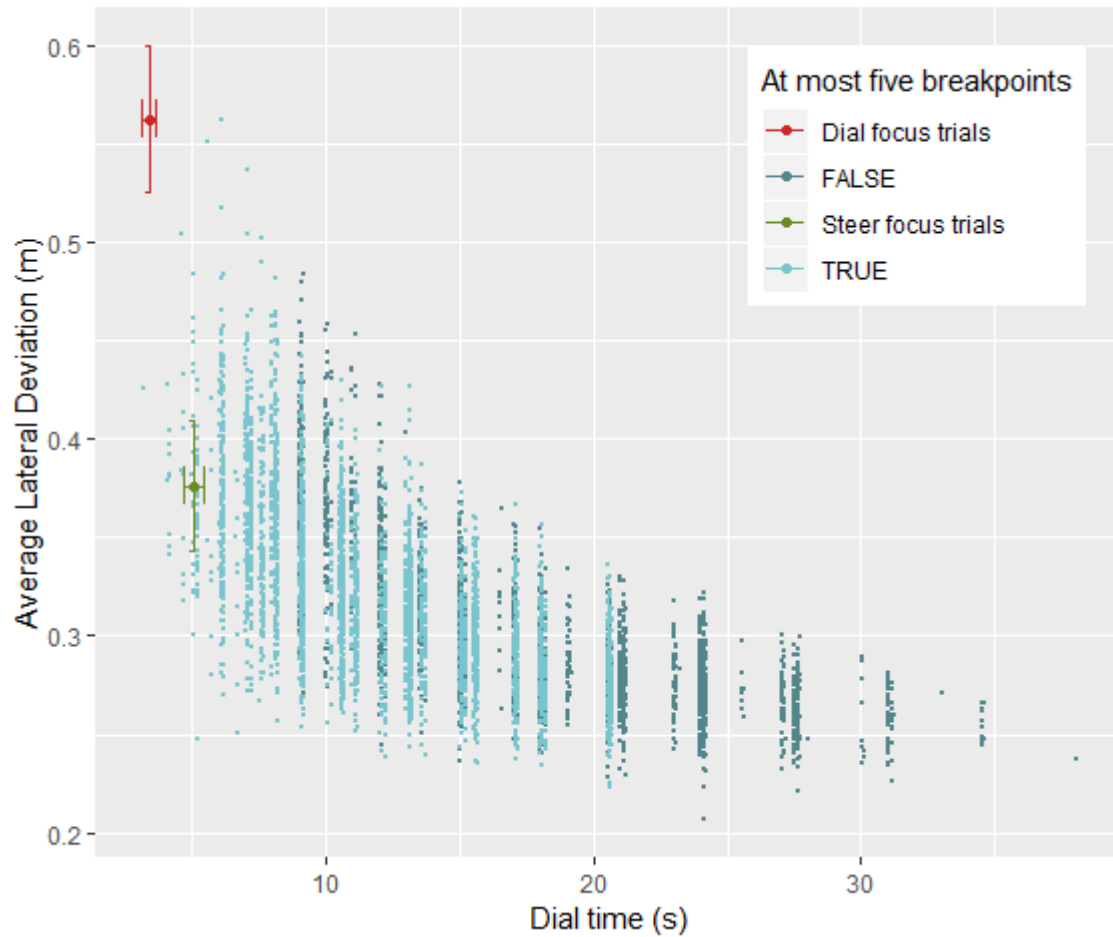
As can be seen in Figure 11, the data from simulations with more than five breakpoints are further away from the human results, and none of them come close to falling within the human range of performance. Based on this, we conclude that it is valid to leave out these strategies. We ran our simulations for different models using all complex strategies where the total number of breakpoints is at most 5.

Note: we only changed one SD value (gaussDeviateSD) when comparing the calibrated drift parameter with the old drift parameter.

Figure 13a: Old drift parameter; old IKI parameter; 10 simulations.
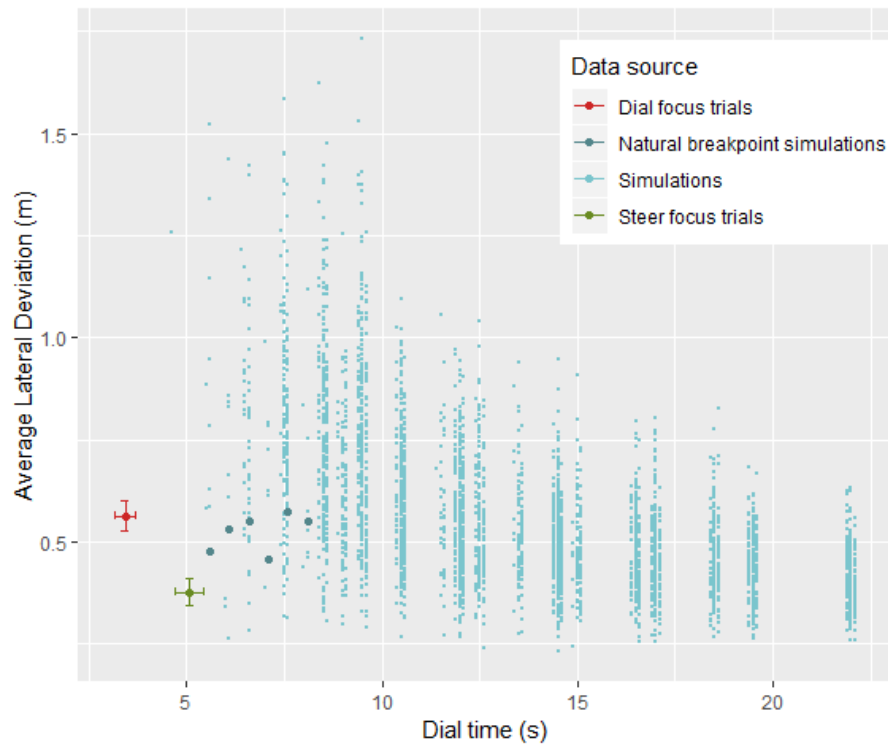

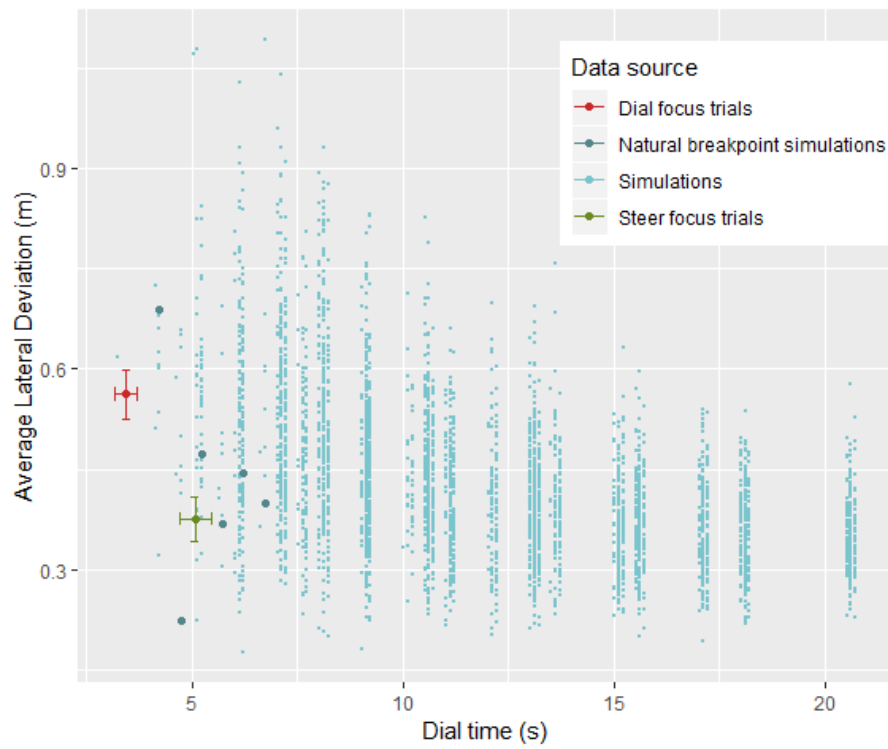Figure 13b: Old drift parameter; new IKI parameter; 10 simulations.

Figure 13c: New drift parameter; old IKI parameter; 10 simulations.
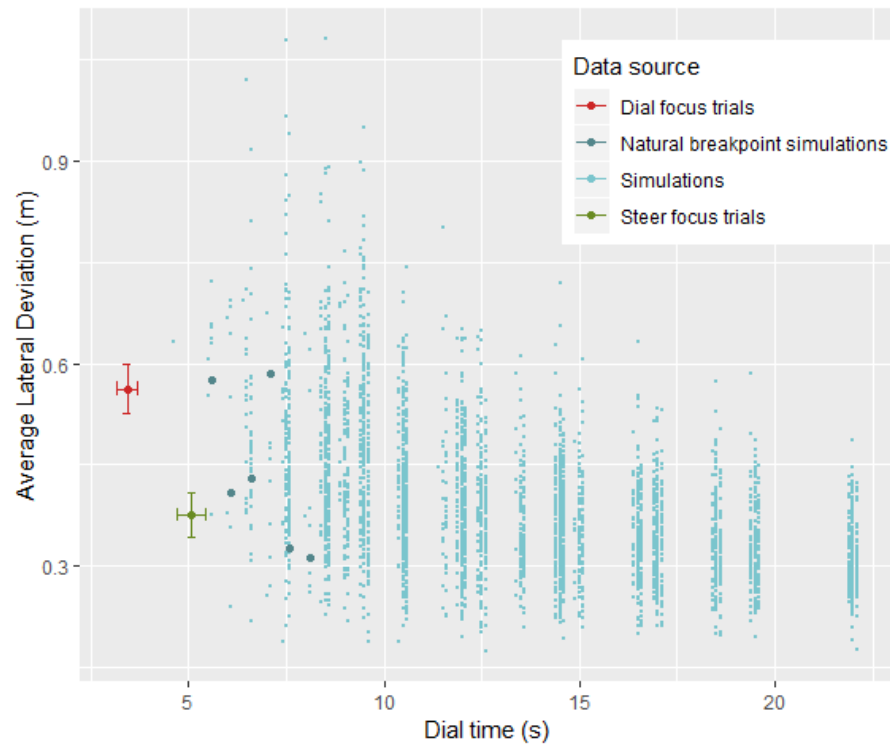

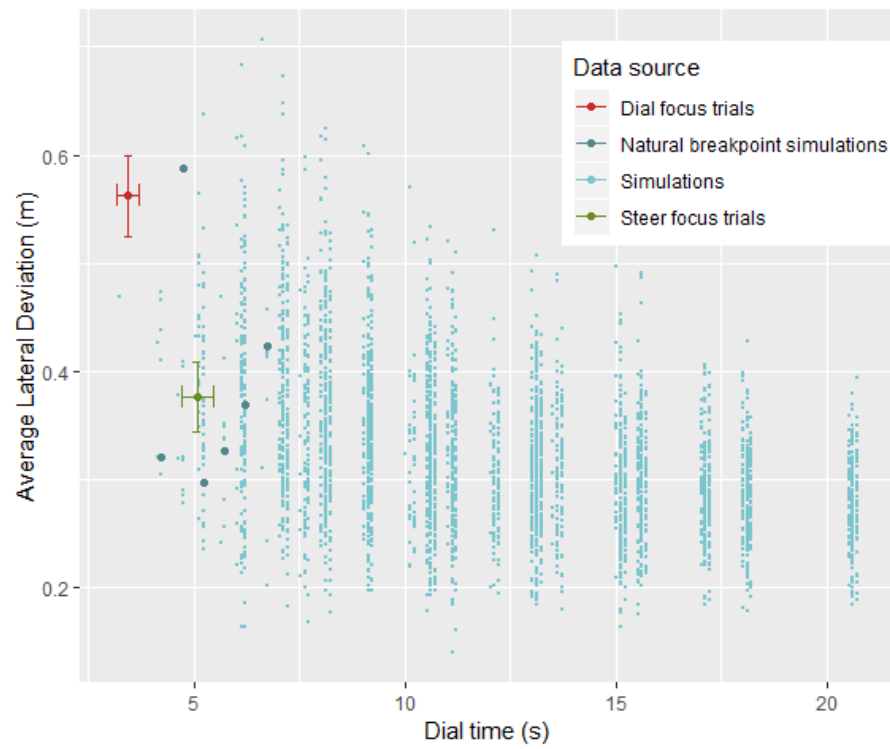Figure 13d: New drift parameter; new IKI parameter; 10 simulations.

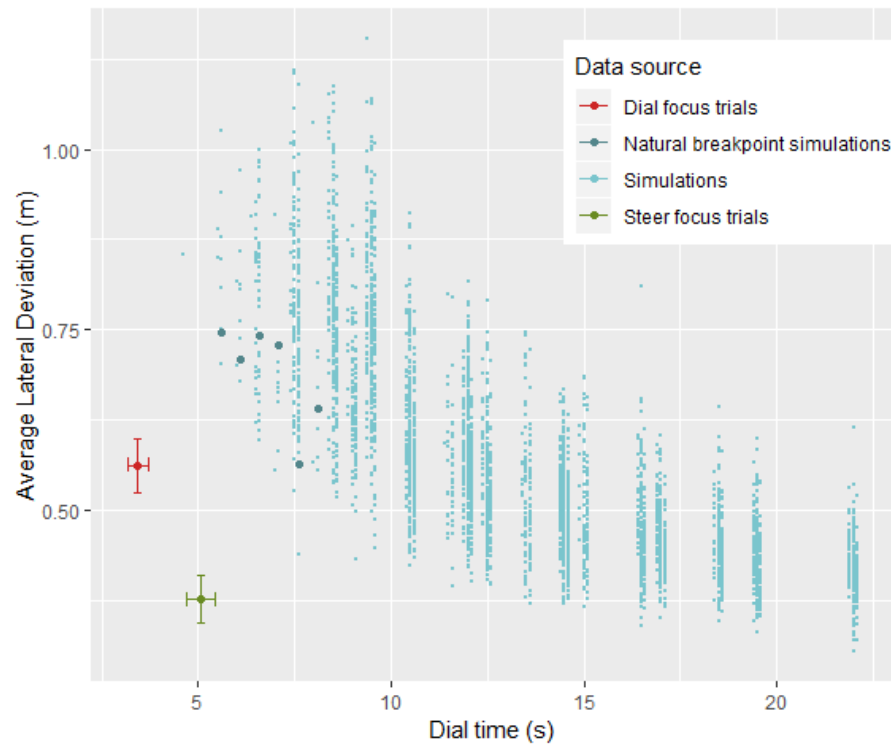Figure 13e: Old drift parameter; old IKI parameter; 50 simulations.

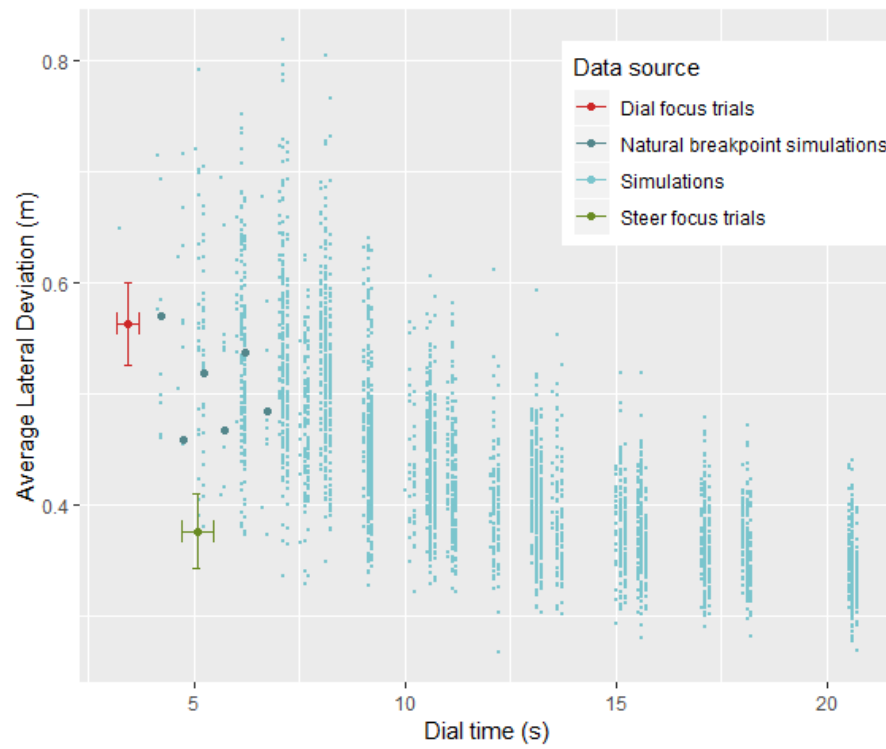Figure 13f: Old drift parameter; new IKI parameter; 50 simulations.

Figure 13g: New drift parameter; old IKI parameter; 50 simulations.
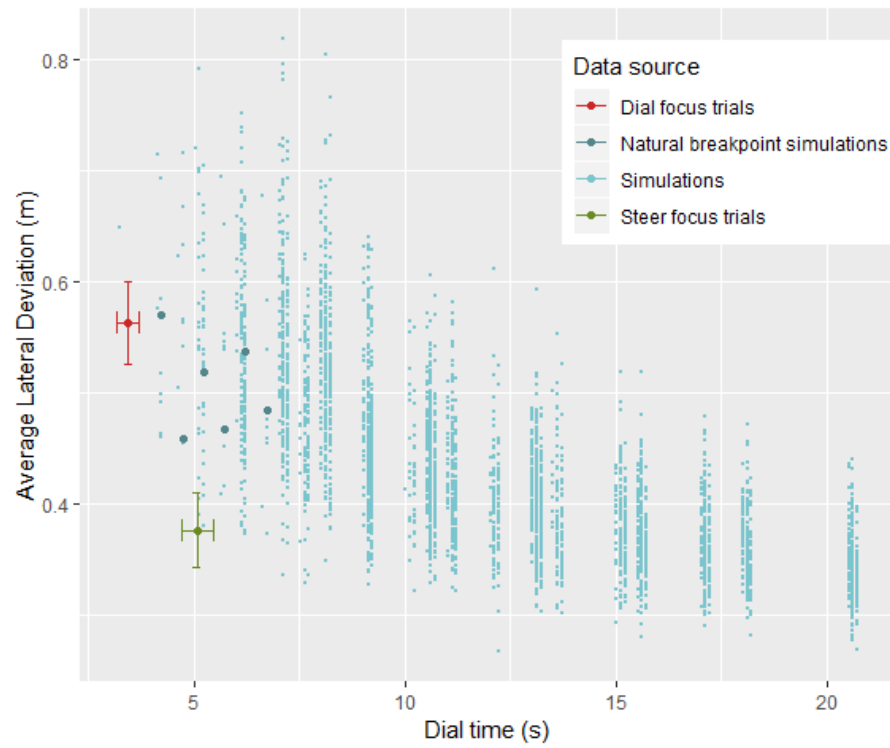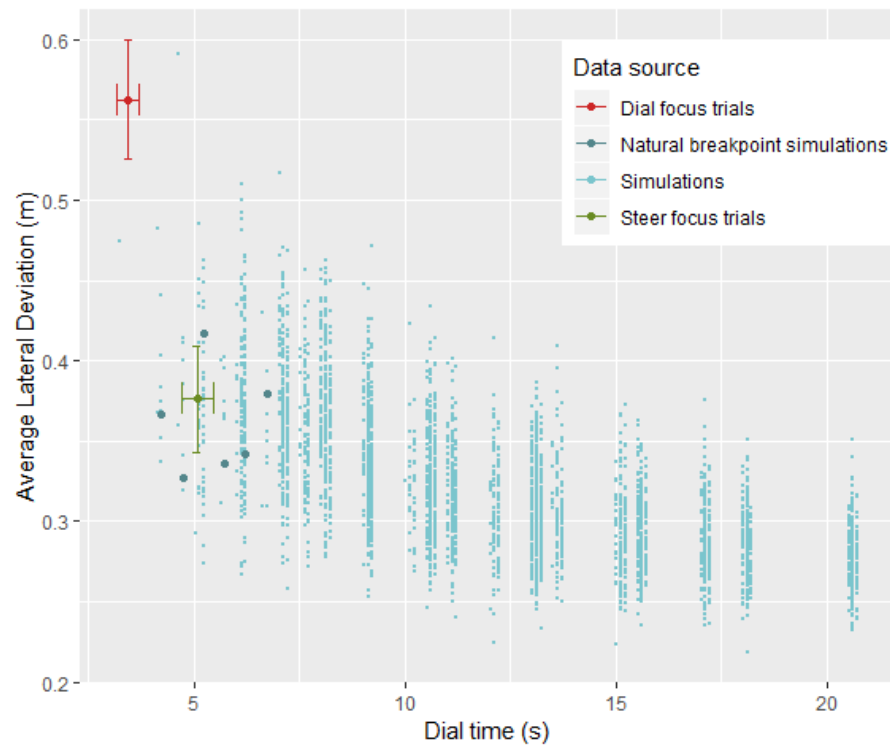

Figure 13h: New drift parameter; new IKI parameter; 50 simulations.

We will consider the model a good fit of the human data if the results for the steering condition fall within the range of model predictions. For a few of the models, the human data fall completely outside of the range of the model; this means that the model does not predict human-like results for any strategy. Since we will use the model to draw a conclusion on what strategies humans use (natural breakpoint or not), we should not use predictions for specific strategies to determine the overall model fit. Therefore, we do not look at the natural breakpoint predictions to determine model fit.

Simulations: if we compare the graphs with 10 simulations (fig. 13a-13d) to those with 50 simulations (fig. 13e-13h), we can see that running 10 simulations seems to lead to a slightly better fit with human data. However, this is due to the spread in lateral deviation being higher (wider range of the y-axis) when there are few simulations, because there is more noise in the data. The overall pattern of the model predictions remains the same irregardless of the number of simulations. As argued by Roberts and Pashler (2000), a good model should sufficiently constrain possible outcomes. The 10 simulation predictions have a better fit only because they restrict the possible outcomes less, by allowing more noise. We conclude that the 10 simulation models are not more valid models than the 50 simulation models.

Drift parameter and IKI parameter: for both parameters we see that by calibrating these parameters the model fit increases compared to their old values. In the model where both parameters have their old values (13e), human data seem to be fairly far off from the model predictions. Then, as one of the parameters is calibrated (fig. 13f, 13g), the model fit increases as the human data fall at the edge of the range of model predictions. Finally, if both parameters are calibrated (fig. 13h) the human data fall well within the range of the model predictions.

For these two parameters, we can add that when calibrating them not only do they produce a better fit, but we have a priori motivation to calibrate their values. Since the experiment used a different simulation and a different dialling interface than the original experiment, it makes sense that these parameters should be updated.

Furthermore, the model seems to be quite reliable/generalizable because by adjusting these two parameters, the model still fits the human data well (fig. 13h), as in the original experiment.

For all 8 models, the human data of the steering condition does not overlap with one of the natural breakpoint strategies. As a result, all models indicate that humans do not seem to interleave solely at the natural breakpoint (fig. 13a-13h). However, this conclusion can only be drawn under the assumption that the model is valid. This at least requires that some other strategy or strategies should fall within the range of human data, which is only true for some models (fig. 13b, 13d, 13f, 13h).