Computational Modeling for Cognitive Science

Portfolio Exam

Ruta Slivkaite

201805872@post.au.dk

School of Communication and Cognition, University of Aarhus,

Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark

Contents

Ruta Slivkaite
201805872@post.au.dk

Assignment 1

Eye-tracking data analysis

https://github.com/PeterThramkrongart/CompMod_1/blob/master/CompMod1.md?fbclid=IwAR39SmGq9GpPhgo3d_qQBVFJSoY8DKLs4Hy2EVpm9UDxbsmPxQsqt_6TXDY

Foraging experiment

Introduction

The conceptual research question, which this study aims to investigate, is how scanpaths during visual search and simple counting tasks differentiate due to both top-down and bottom-up factors. Specifically, we will focus on the differences in the saccade lengths and their distributions in the two conditions - count and search task.

The impact of this question is more conceptual. We are interested in uncovering how task goals can influence an otherwise quite bottom-up driven process.

Hypothesis

We hypothesize that the search or *foraging* condition will differ systematically from the count condition. In the foraging condition, we expect to observe a heavy-tailed distribution of saccade lengths and that the saccade amplitudes will be longer, compared to the count condition. This would imply that bottom-up features of the stimulus are not the only factor influencing the eye gaze movements, but that the top-down factors such as the type of task can impact the pattern of eye movements as well.

Methods

Participants

There were 6 participants in total in the foraging experiment, all of them 4th-semester Cognitive Science students at Aarhus University. One goal during the data collection was to introduce confounding factors that would intervene with the eye tracker's ability to detect the participants'

 Ruta Slivkaite
201805872@post.au.dk

pupils. Hence, most participants were wearing contact lenses or glasses and/or a lot of mascara on their eyelashes. This handicap was added so that we can observe what 'real', messy data looks like and have the opportunity to learn not only how, but also why, it is necessary to control for these confounds in future experiments. Furthermore, during one of the trials the research assistant accidentally exited the PsychoPy script, resulting in loss of data.

Task

The experiment was comprised of two tasks:

(1) Search condition: in this condition, participants viewed images of natural scenes for a set amount of time, with the goal of finding a transparent star, hidden in the picture. If the participant located the star, they were instructed to fixate on it.

(2) Count condition: in this condition, participants viewed images of natural scenes for a set amount of time, in order to count elements depicted in the image (e.g. counting the number of birds or sheep). The element which the participant was supposed to count was not explicitly stated, but rather inferred from features of the image itself.

Eye-Tracker Setup

The participants' gazes were measured with the EyeLink1000 at COBE Lab. The machine has a monocular measure, which means that we need to measure the participant's dominant eye, which can be tested by the Miles test. After setting the sample rate at 500 Hz and illuminator power at 75%, calibration of the eye ensued, in order to make sure that the eye tracker could detect the pupil from different angles as accurately as possible.

Preprocessing Eye-Tracking

Using the EyeLink1000 software, the visual eye-tracking data was translated and grouped into saccades and fixations of differing lengths. Consequently, the scan paths were determined (*Figure 1*). This data was then exported from the eye-tracker. Afterwards the data was cleaned by deleting a few nonsensical outliers, which were deemed artifacts from the eye-tracker.

Ruta Slivkaite
201805872@post.au.dk

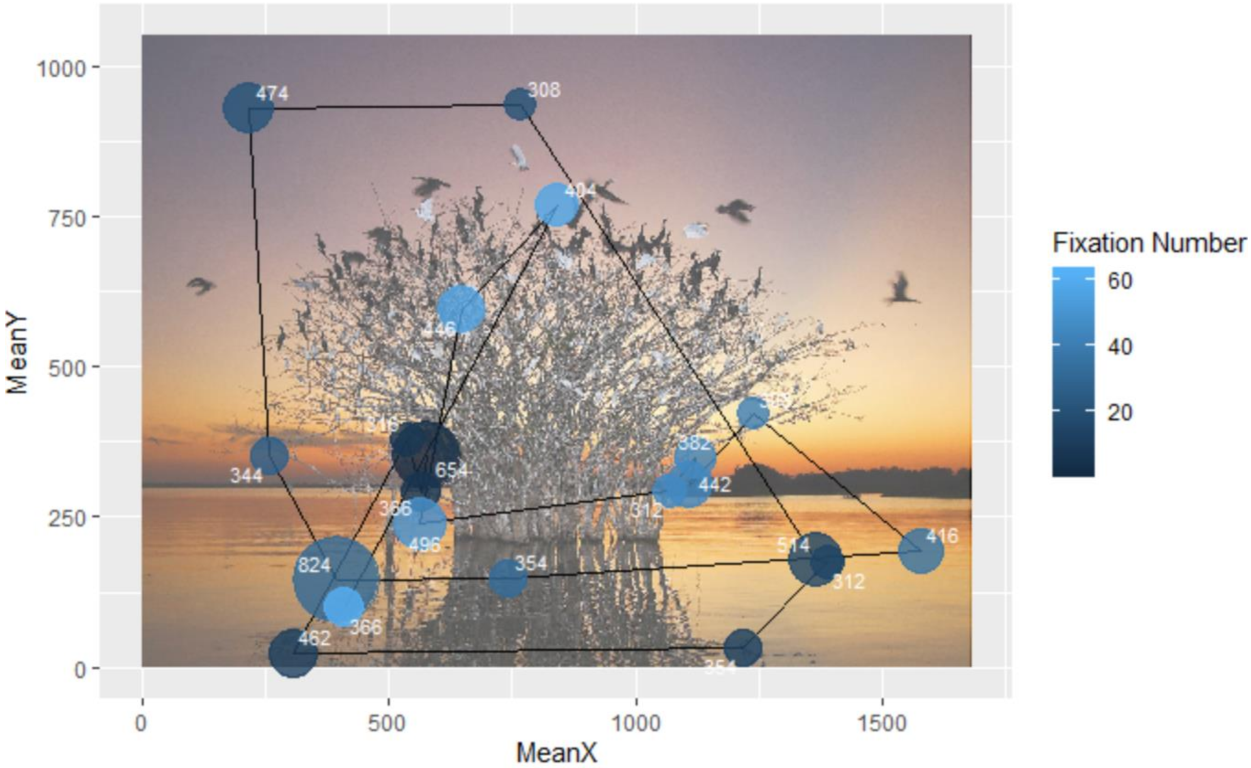


Figure 1, scanpath example with saccades and fixations

Statistical Modeling

We fitted a linear mixed-effects model (LMM) to the data, where saccade amplitude is predicted by the fixed effect condition (count/search), while taking the random effects of participants and stimuli into account. Further, these random effects of ID's and stimuli may vary according to condition.

Saccade Amplitude ~ 1 + Condition + (Condition|ID) + (Condition|Stimulus)

After a visual inspection of the model residuals via a distribution plot, we find the data to be non-normally distributed. In particular, the distribution has long positive tails and is skewed towards zero, since there are naturally no negative values of saccade amplitudes. In an attempt to combat this and maintain our assumption of normally distributed data points, we decide to transform the saccade amplitude to log scale.

However, after comparing residuals of log scale and gaussian scale model with a simulation-based approach (DHARMa) (see *figure 2* and *figure 3*), we discovered two things. Firstly, the

Ruta Slivkaite
201805872@post.au.dk

transformation did not remedy the non-normalcy of the data, secondly, both models have low Goodness of Fit (GOF). Despite this, we continue the analysis using the log-transformed data, because this is the method we often utilize, when we have non-normal positive, heavy-tailed distributions.

In order to transform the coefficients of the model (log-odds) back to the correct scale, informing us about the angle sizes of the visual field and thus amplitude of the saccades, we exponentiate the beta values.

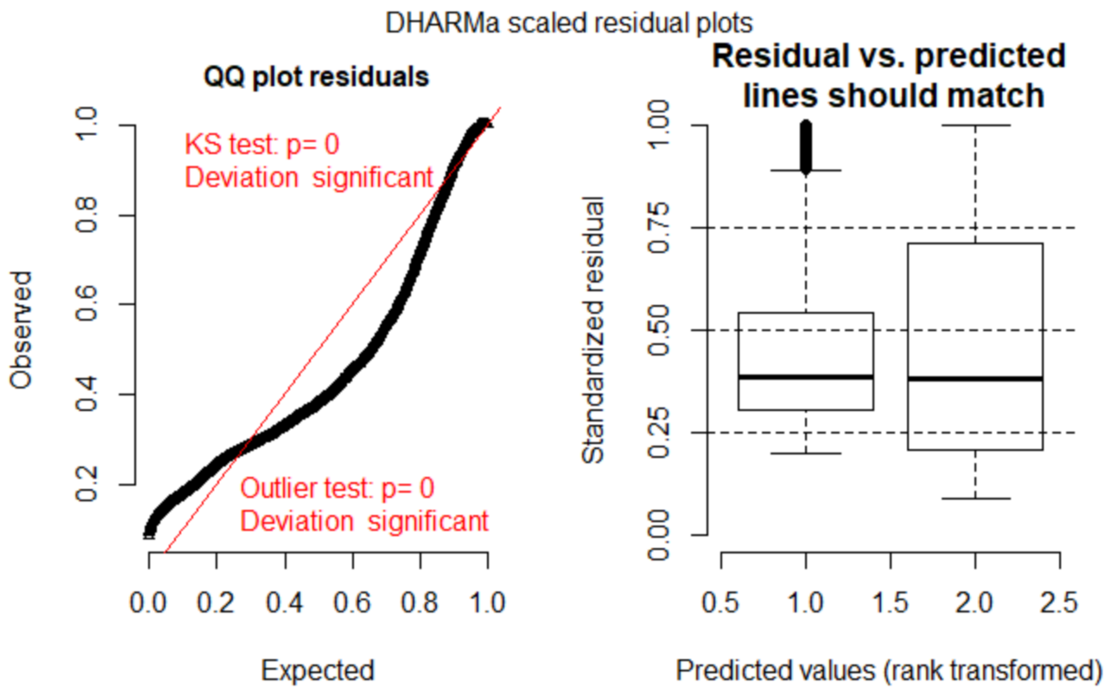


Figure 2 - residual plots from gaussian model

Ruta Slivkaite
201805872@post.au.dk

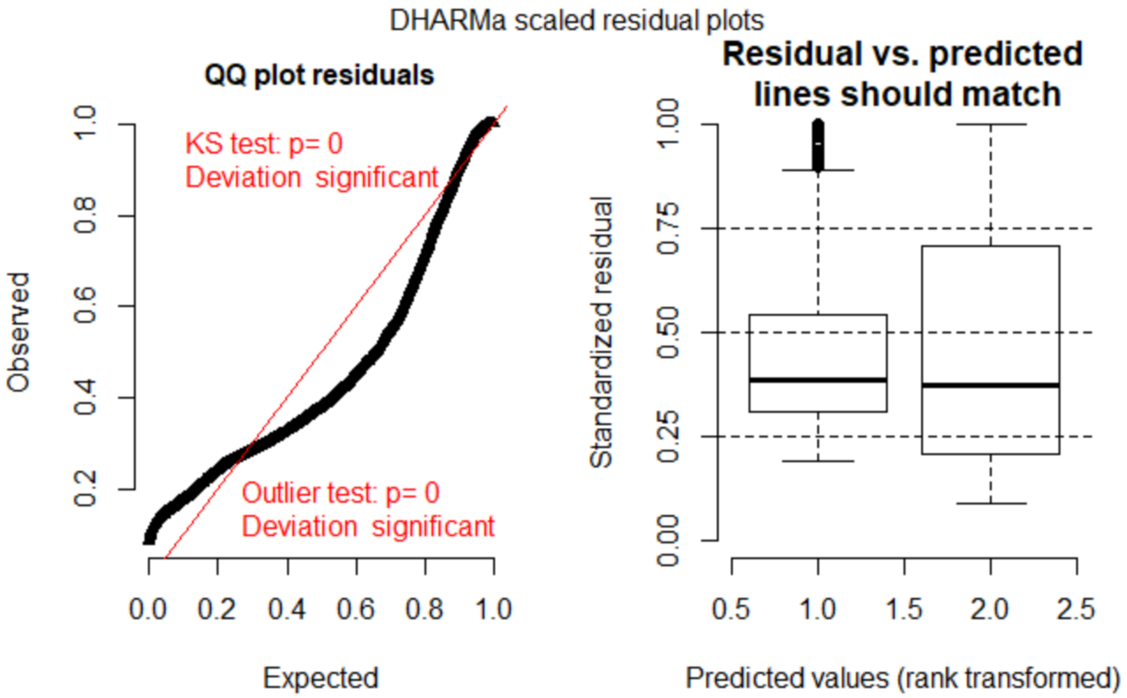


Figure 3 - residual plots from log-transformed model

Results

The results of the mixed effects analysis is depicted in *table 1* (on the log scale). According to the model the average amplitude in the foraging condition is larger by 1.8 degrees compared to the count condition.

Condition	Beta-value	SE	p-value
Search	0.53761	0.07907	*<.001*

Table 1 - results from model analysis

We also calculate the marginal R^2 of the model, which is 0.00529. This signifies that the model can account for 0.529 % of the variance.

In terms of whether or not the tasks will yield a data distribution with heavy tails, there seems to be a visual difference between the two conditions. In *figure 4*, we see a density distribution over the saccade amplitudes. We observe heavy-tails in both conditions, but the distribution in the

Ruta Slivkaite
201805872@post.au.dk

foraging or search condition is more distinct. However, a formal analysis, such as the Hill Estimator, is needed to confidently draw a conclusion about the significance of this difference.

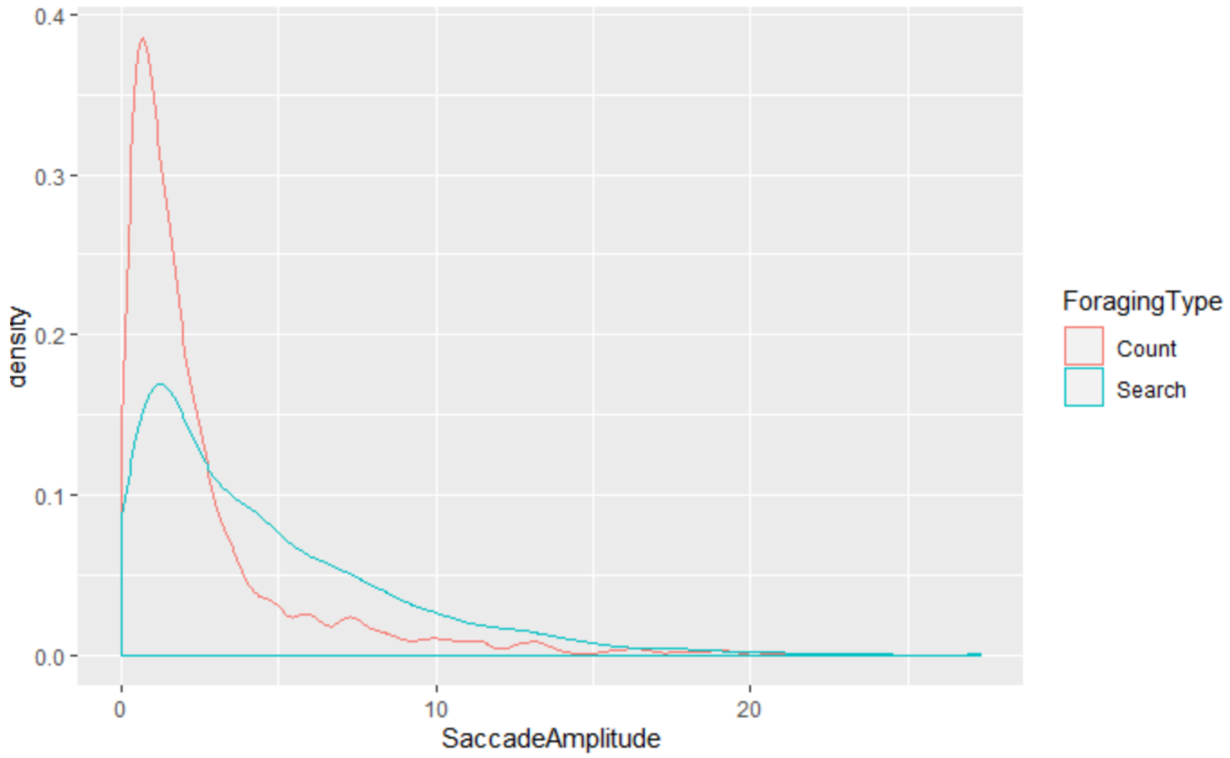


Figure 4 - density distribution of saccade amplitudes

Discussion

The results of the model analysis revealed that there is a significant difference between the duration of the saccades in foraging type tasks and counting tasks. As hypothesized, we found a larger average amplitude during foraging compared to the count condition, meaning that the foraging task elicits a response in the participants, which influences their scanpaths between fixations to become longer. This result is congruent with our hypothesis and implies that the task at hand makes a difference to how participants move their gazes, specifically that the saccade amplitudes are lengthened during visual foraging. See *figure 5*, which depicts the linear model superimposed on the distribution of saccade amplitudes between conditions. The search condition distribution has noticeably heavier tails compared to the count condition.

The results imply that condition and therefore top-down processes of the brain influence how visual tasks perform in modulating the saccade amplitudes. Nonetheless, there are several problems in the experiment, which need to be addressed. First off, having only 6 participants always imposes a great constraint on the generalizability of the results. Second, the simulation-based approach of

Ruta Slivkaite
201805872@post.au.dk

the model's residual evaluation suggests that overall model performance is quite low. Therefore, it is good practice to be rather guarded when concluding on the results.

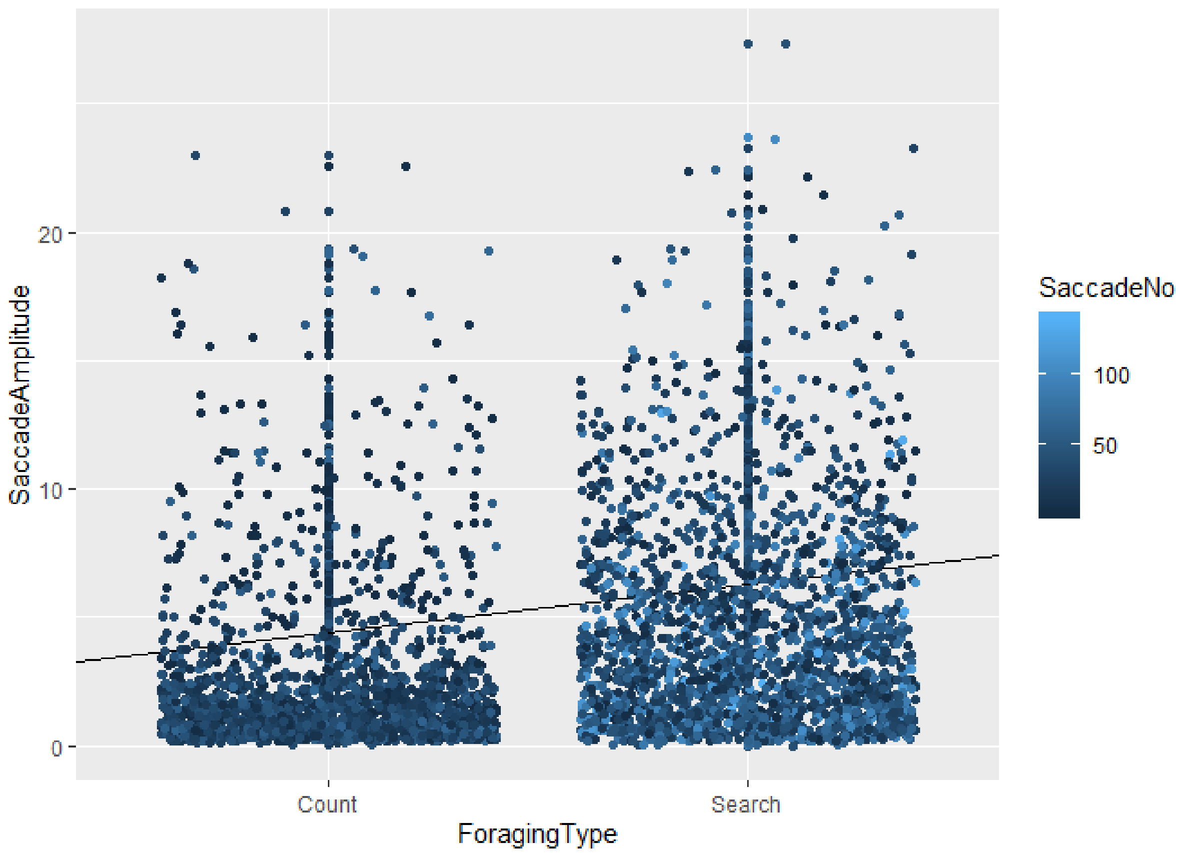


Figure 5 - saccade amplitude divided by condition with regression line

Social engagement experiment

Introduction

Ruta Slivkaite
201805872@post.au.dk

The conceptual research question of this study was to investigate how social cognition differs during social interaction compared to social observation. More specifically, we want to elucidate whether there is a difference in physiological arousal, determined by pupil size, depending on the social situation. We look at whether two types of ostensive cues used to simulate different states of social engagement, affect pupil size. In other words, does social engagement versus observation matter in how we process social information?

These research questions have implications for research on social cognition in everyday life and in clinical conditions. For instance, if we were able to distinguish social interaction from social observation, we would have a good argument to postulate that there might, in fact, be two separate models of social cognition. This could be employed in detecting impairments related to mental or developmental disorders that show deficits in social interactions like people on the autism spectrum.

Hypothesis

We hypothesize that interactive, engaging stimuli would elicit more emotional arousal, resulting in a greater pupil dilation than stimuli affording a more observational attitude in the participant. In other words, the condition relying on two ostensive cues rather than 1 or 2 would cause the greatest physiological arousal compared to other conditions (see the *Methods* section) .Furthermore, we predict that there might be an interaction effect between the cues. We hypothesize that two ostensive cues in concert will elicit greater physical arousal than the cues on their own.

Methods

Participants

There were 6 participants in total in the social engagement experiment, all of which were 4th-semester Cognitive Science students at Aarhus University. One of the goals during data collection was to introduce confounding factors that would intervene with the eye tracker's ability to detect the participants' pupils effectively. Hence, most participants were wearing contact lenses or glasses and/or a lot of mascara on their eyelashes. This handicap was introduced to prompt our work with 'real', messy data.

Task

The task directly compared interactive and observational social cognition in order to test the hypothesis. Participants watched short videos with either a male or female actor. The actor in the

Ruta Slivkaite
201805872@post.au.dk

video placed an object on the table performing either one, two, or no social cues. The different video clips therefore varied in their level of social engagement, enabling us to test the hypothesis.

The experiment has four conditions comprised of a two times two factorial design;

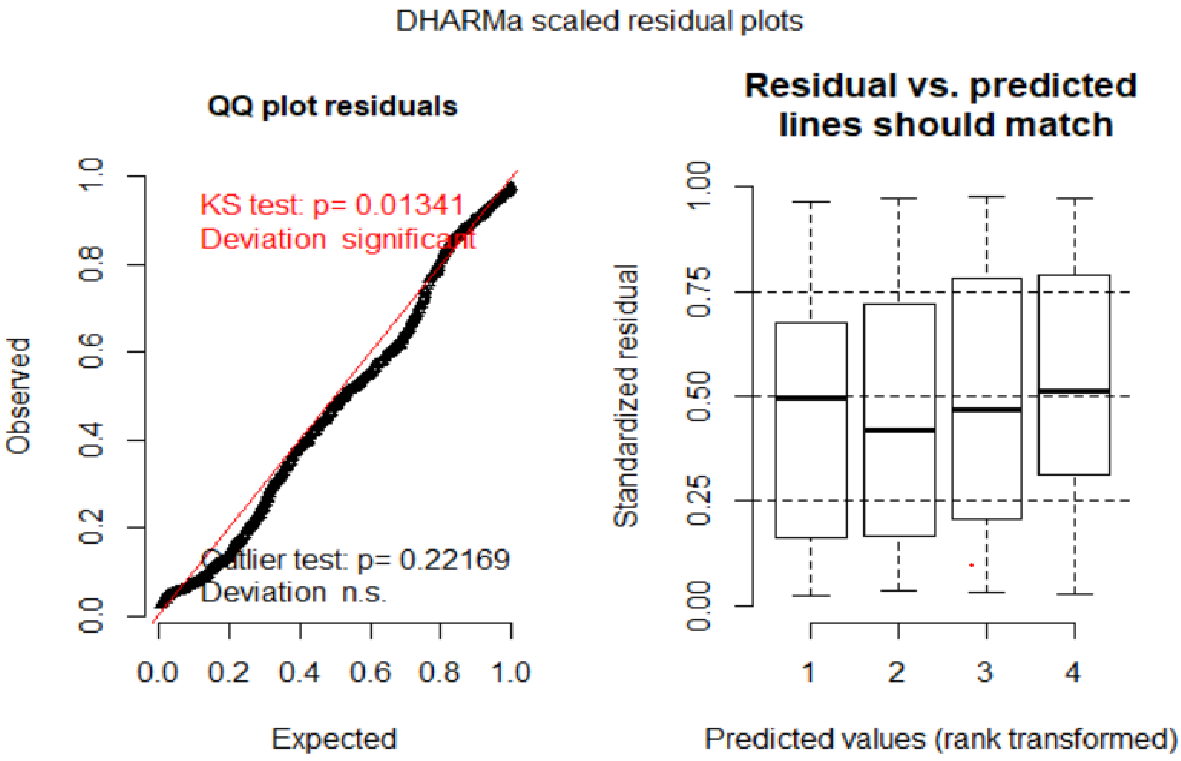
(1) ostensive and direct: An actor facing the direction of the participant (direct), exhibits an interaction-initiating cue by raising their eyebrow (ostensive) before performing the object-related gesture.

(2) non-ostensive and direct: An actor is facing participan (direct), but does not exhibit ostensive cues (non-ostensive)

(3) ostensive and averted: An actor sits at an averted angle, facing someone outside the perspective of the camera (averted) and exhibits the eye-brow lift and nod (ostensive).

(4) non-ostensive and averted: An actor sits at an averted angle, facing someone outside the perspective of the camera (averted) and doesn´t exhibit ostensive cues (non-ostensive).

Shifting between the conditions with ostensive cues vs no ostensive cues and altering the direction of the actor would shift the experience of participants between observational mode and an interactive mode. The most interactive condition is (1), while the most observant condition is (4). This enables us to ask the question whether social observation is any different from social interaction..



Ruta Slivkaite
201805872@post.au.dk

Figure 6 - residual plots from gaussian model Eye-Tracker Setup

The participants' gazes were measured with the EyeLink1000 at COBE Lab. The machine has a monocular measure, which means that we need to measure the participant's dominant eye, which can be tested by the Miles test. After setting the sample rate at 500 Hz and illuminator power at 75%, calibration of the eye ensued, in order to make sure that the eye tracker could detect the pupil from different angles as accurately as possible

Preprocessing Eye-Tracking

Using the EyeLink1000 software, the visual eye-tracking data was translated and grouped into saccades and fixations of differing lengths. Consequently, the scan paths were determined. This data was then exported from the eye-tracker. Afterwards the data was cleaned by deleting a few nonsensical outliers, which were deemed artifacts from the eye-tracker.

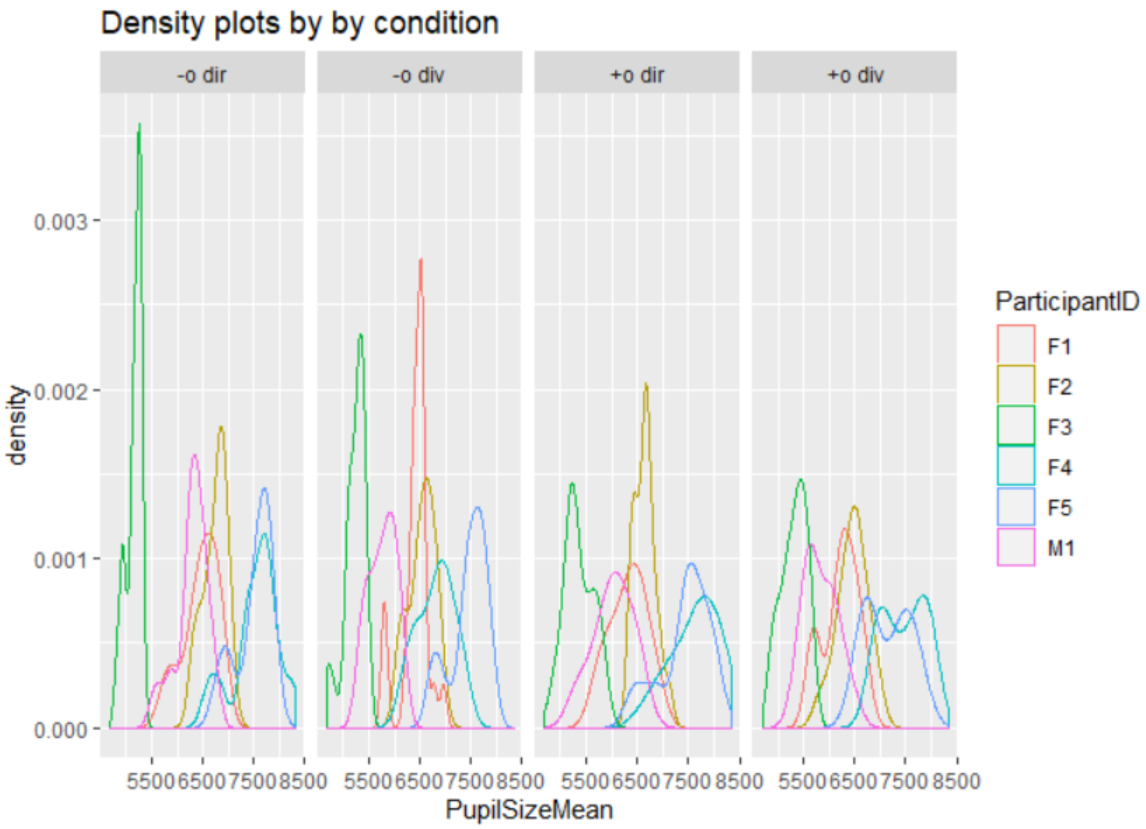


Figure 7 - density plot for each participant

Statistical Modeling

Ruta Slivkaite
201805872@post.au.dk

We used two-by-two factorial design (making up four conditions) with the main factors Direction and Ostension, each of them having two levels.

We fitted a linear mixed-effects model (LMM) to the data, where pupil size is predicted by the fixed effect Direction (direct/averted) and its interaction with the fixed effect Ostension (ostensive/non-ostensive), while taking the random effects of participants into account. Each participant is expected to have individual intercepts and slopes dependent on the interaction between Direction and Ostension.

Pupil Size Mean ~ 1 + Direction * Ostension + (1 + Direction * Ostension | Participant ID)

We were able to assume normalcy of the data after looking at the model residuals using DHARMA (see *figure 6*) as well as the density plots in *figure 7*.

Results

The output of the model is reported in *table 2*. See *figure 8*, which depicts the distribution of the pupil size by condition in a box plot. Visually, it is already evident that the directed condition without ostension affords the largest pupil sizes, while the averted (div) condition without ostension seems to be correlated with smaller pupil sizes.

We do not report the R^2 because there are no significant effects to analyze,

Condition	Beta-value	SE	p-value
Direct	250.29	128.011	0.108
Ostension	45.65	138.480	0.755
Direct:Ostension	-92.69	131.254	0.511

Table 2 - results from model analysis

Ruta Slivkaite
201805872@post.au.dk

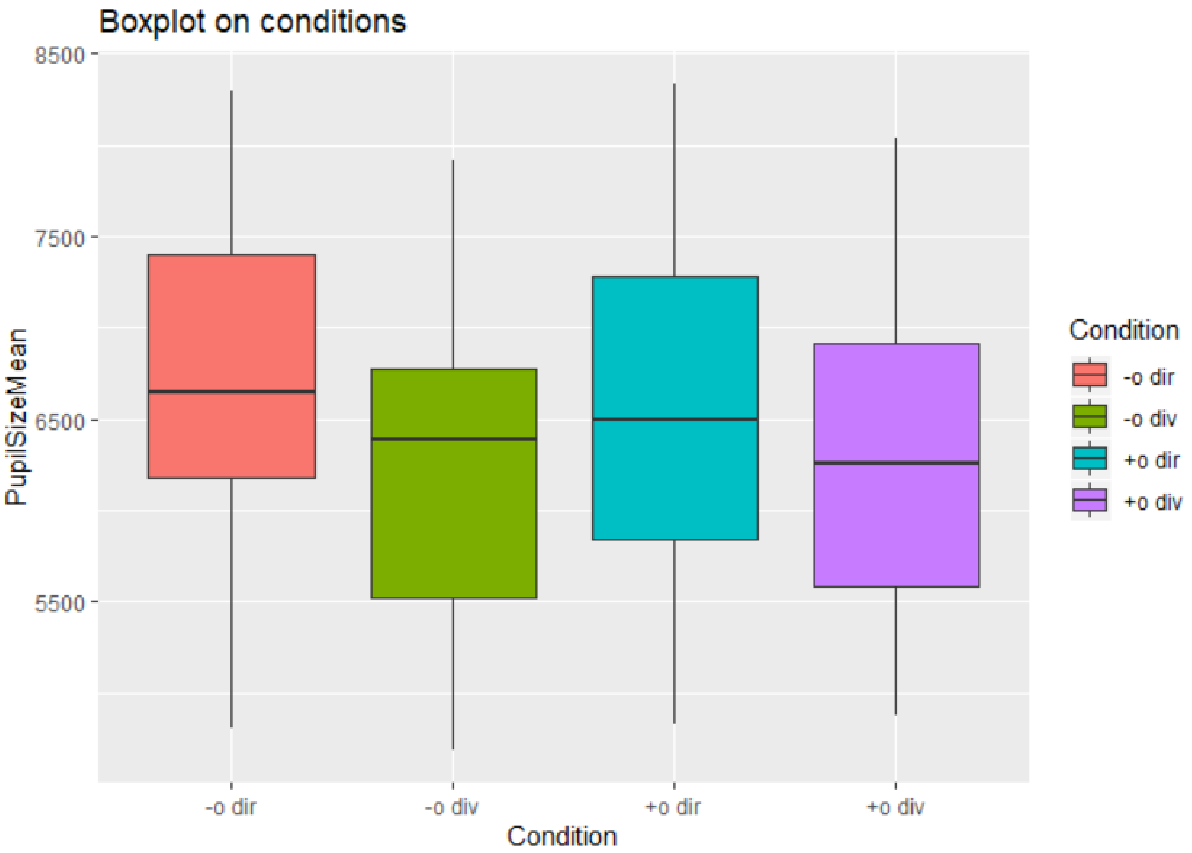


Figure 8 - boxplot over pupil size in the four conditions

Discussion

The results from our model imply no significant difference between the different social situations. According to the model, pupil size and therefore physical arousal increases, when the interlocutor is directed towards the participant compared to when they are facing in a different direction.

Further, pupil size is not significantly dependent on the interaction effect between direction and ostension. According to the results pupil size shrinks when the actor, who is directed towards the participant, raises their eyebrows in an ostensive gesture. On the other hand, when the actor is facing away from the participant, the ostensive cue causes the pupil size to increase, implying arousal. See *figure 9*, depicting the interaction effect, which, however, is non-significant.

Against the background of these results, we can conclude that the pupil size is not significantly different in the social observation and social interaction scenario. Our results do, therefore, not confirm our hypotheses. We also predicted to observe the most arousal when both social cues (Direction and Ostension) were present, but this was not the case. We found that the ostensive cue actually shrunk the pupil size when coupled with the direct condition, even though these results were, again, not significant.

Ruta Slivkaite
201805872@post.au.dk

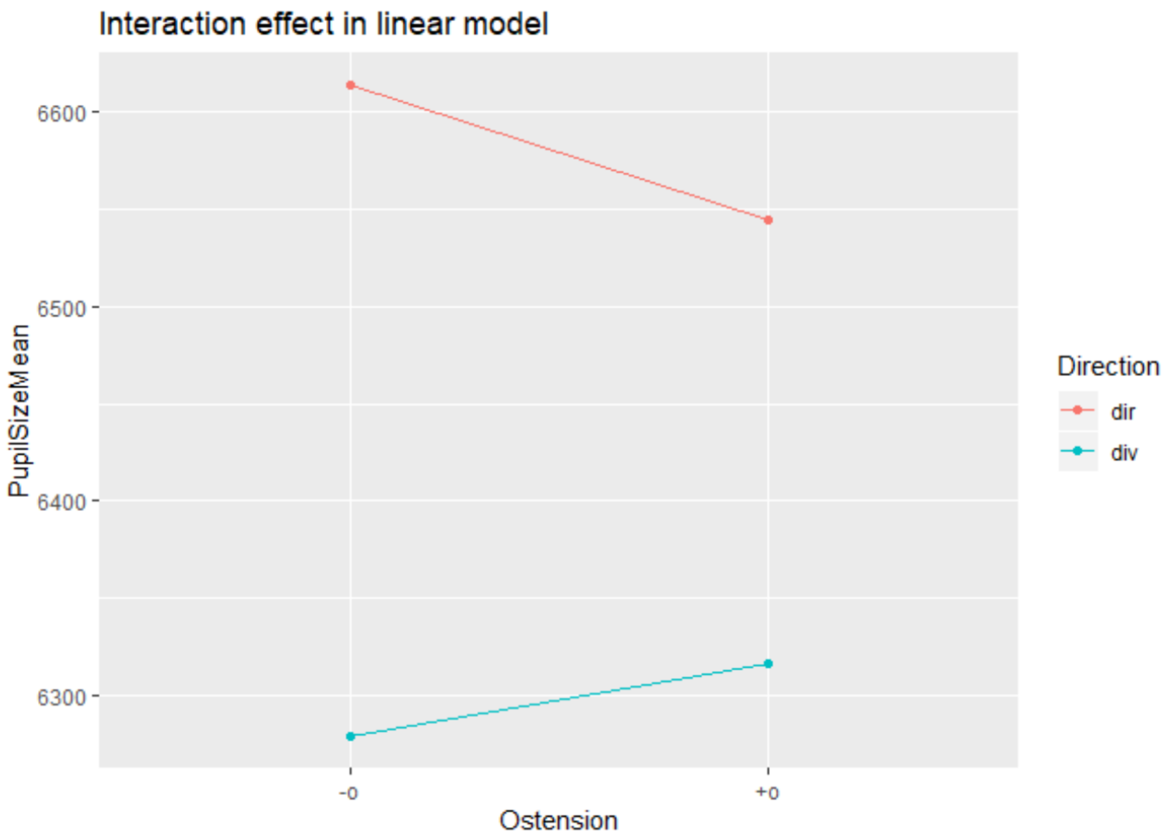


Figure 9 - plot of the interaction effect between ostension and direction

Package Reference

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL: http://www.rstudio.com/.

Hadley Wickham (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://CRAN.R-project.org/package=ggplot2.

Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48.<doi:10.18637/jss.v067.i01>.

Kamil Slowikowski (2019). ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. R package version 0.8.1.: https://CRAN.R-project.org/package=ggrepel

Ruta Slivkaite
201805872@post.au.dk

Florian Hartig (2020). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.2.7.: https://CRAN.R-project.org/package=DHARMa

Ruta Slivkaite
201805872@post.au.dk

Assignment 2

Evaluating the Cognitive science knowledge of Cognitive science teachers

Github: https://github.com/PeterThramkrongart/CompMod2/blob/master/assignment-2-last-document.md

This assignment centres around evaluating rates from a binomial distribution, using the case of assessing the teachers' knowledge of Cognitive Science. In order to execute these assessments we utilize bayesian statistics. The key intuitions we gained from this assignment was how priors and sample sizes influence the posterior distribution in interactive ways.

PART ONE

1. Riccardo's estimated knowledge of CogSci using a uniform prior, Uniform(0,1)

Riccardo estimated knowledge of cognitive science is normally distributed with a MAP 0.5 (SD = 0.2) which is depicted in the plot below (figure 1) and was computed using a grid approximation.

Using a quadratic approximation the computed posterior distribution has the mean 0.5 (SD = 0.2). This is the same result we get using a grid approximation, because the MAP and the mean are the same value whenever we have a normal distribution.

According to our calculations, the probability that Riccardo knows more than chance, given a flat prior and a correct answering rate of 3:6, is 50 %. This is due to 50 % of the probability distribution being > 0.5 (see figure 1).

Ruta Slivkaite
201805872@post.au.dk

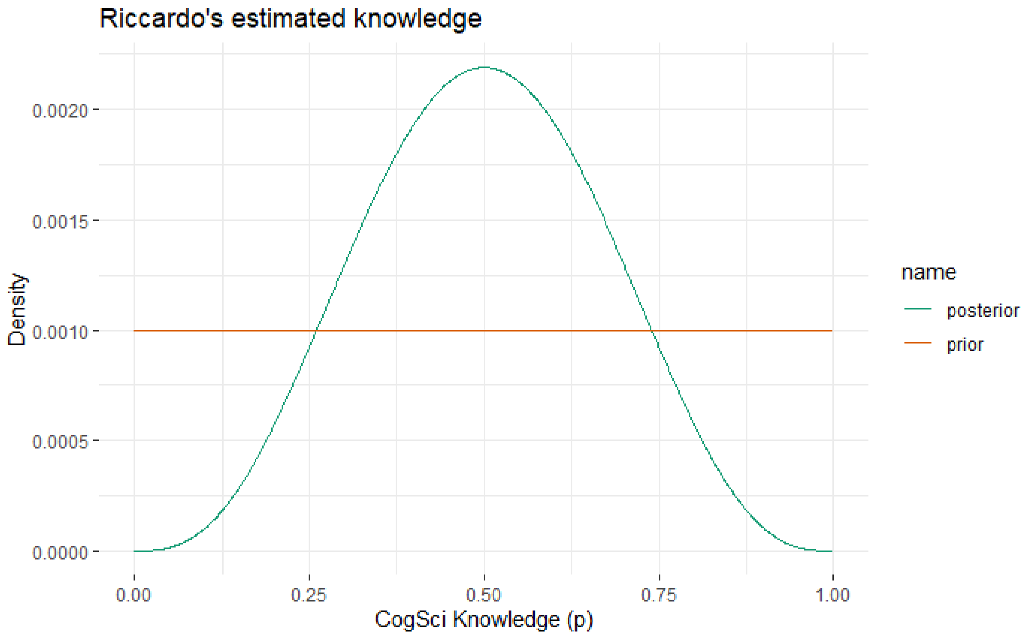


Figure 1 - Illustration of the Riccardo´s estimated knowledge with the uniform prior

2. Estimation of all the teachers' knowledge of CogSci using grid approximation and the uniform prior, uniform(0,1)

Teacher	Lower HPDI of 80 %	Upper HPDI of 80 %	MAP	SD
Riccardo	0.26	0.70	0.50	0.17
Kristian	0.58	1	1	0.20
Josh	0.77	0.84	0.81	0.28
Mikkel	0.45	0.56	0.5	0.04

Table 1- Teachers' Cognitive Science knowledge estimated by grid approximation

Ruta Slivkaite
201805872@post.au.dk

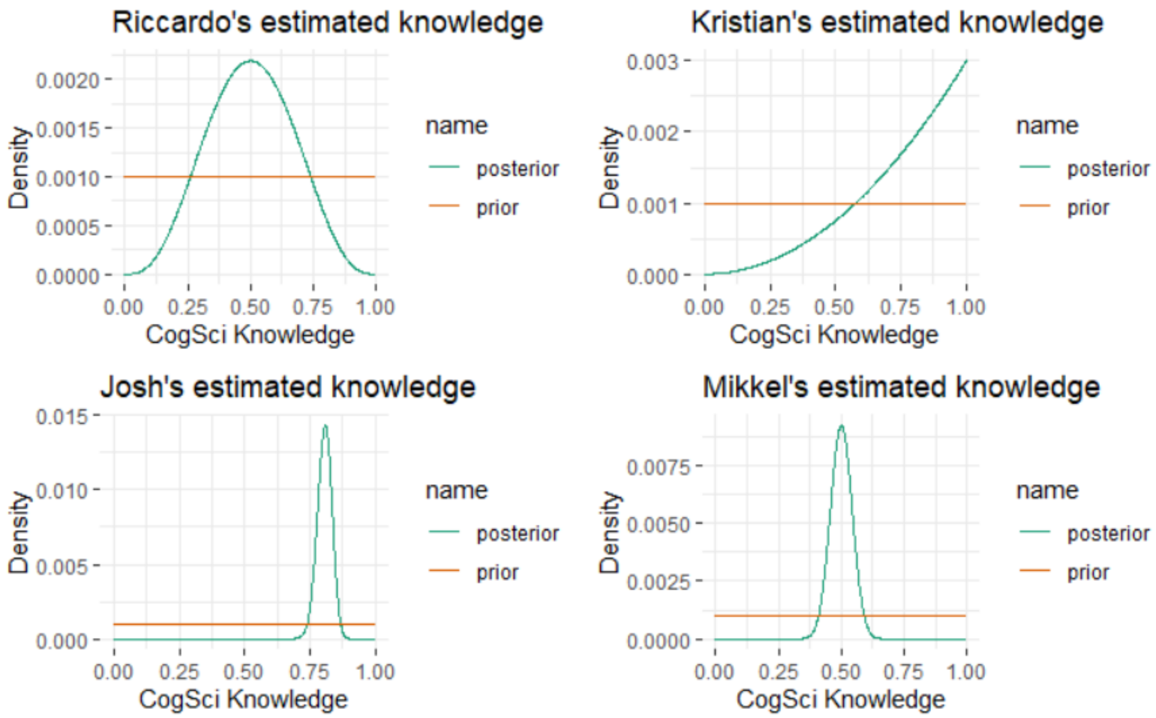


Figure 2 - Illustration of the uniform prior and calculated posterior of teachers' knowledge

Josh's narrow and relatively high HPDI (0.77, 0.84) and high MAP (0.81) reflects that he got many questions correct out of a large number. Compared to Kristian who, despite having the highest MAP value (1), has very broad HPDI (0.58, 1) due to very few answered questions. We can see that while MAP is primarily affected by how many questions were correctly answered, HPDI takes into account as well what is the overall number of questions and thereby gives us more information.

On the other hand, Mikkel and Riccardo have similar MAP estimates, but the spread is very different (see figure 2). Mikkel (MAP = 0.5, SD = 0.04) has answered more questions and therefore his estimates are more certain than Riccardo's (MAP = 0.5, SD = 0.17). Whether Riccardo or Mikkel is the most proficient in cognitive science is hard to say. The model is more certain that Mikkel knows *something* about cognitive science, but Riccardo has the opportunity to surprise more - both in a positive and negative direction.

When determining who is the smartest teacher, we can exclude Riccardo and Mikkel immediately as their performances are about chance level. To compare Joshua and Kristian, we contrasted how much the posterior probability of Kristian's knowledge lies below the posterior probability of Joshua's knowledge. Using this method we calculated that Joshua has a 51.35 % chance of being smarter than Kristian which almost equals to chance. Furthermore, there is too much uncertainty due to Kristian's low sample size to confidently say which teacher is the most knowledgeable.

Ruta Slivkaite
201805872@post.au.dk

3. Estimation of teachers' knowledge given an informative prior - Normal(0.8, 0.2)

Teacher	Lower HPDI of 80 %	Upper HPDI of 80 %	MAP	SD
Riccardo	0.48	0.80	0.65	0.13
Kristian	0.69	1	0.89	0.13
Josh	0.77	0.84	0.80	0.03
Mikkel	0.45	0.56	0.51	0.04

Table 2 - Teachers' knowledge estimated with informed priors

The plots (displayed in figure 3) illustrate the prior and posterior distribution of each teacher given the informative prior. These plots and table 2 show that the more informed prior has changed teachers´ posteriors in a way so that they were shifted to the right. This implies that the assumption (prior) that our teachers are pretty smart increases the probability of their predicted knowledge being higher.

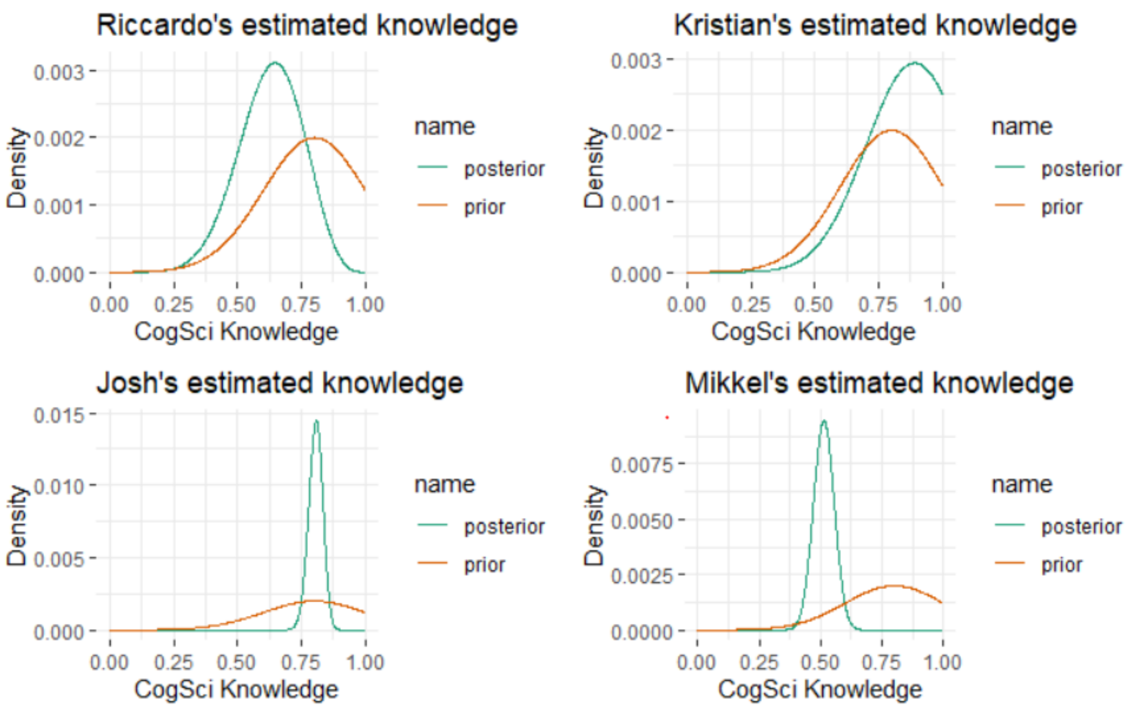


Figure 3 - Illustration of prior and posterior distribution of teachers' knowledge

Ruta Slivkaite
201805872@post.au.dk

4. Estimation of teachers' knowledge given a larger data set

If we multiply all teachers' answers with 100, we get new values for some teachers while other estimates remain close to constant.

For the new priors, if the teacher's initial sample size was small, multiplying their sample size with 100 makes the HPDI more narrow. This makes perfect sense since the model becomes more certain. Moreover, more weight will be given to the observed data, reflected in the MAP which is pulled further in the direction of the observed data compared to the prior. We don't see a big difference between uniform and informed priors since the priors have been "overruled" by the sheer amount of data.

This means that Kristian is suddenly more likely to be the most proficient teacher in cognitive science knowledge. It also implies that when the updating sample size becomes larger, the prior becomes less and less influential.

Teacher	Lower HPDI of 80 %	Upper HPDI of 80 %	MAP	SD
Riccardo	0.47	0.52	0.50	0.02
Kristian	0.99	1	1	0.005
Josh	0.81	0.81	0.81	0.003
Mikkel	0.49	0.50	0.50	0.004

Table 3 - Estimates of teachers' knowledge using the uniform prior

Teacher	Lower HPDI of 80 %	Upper HPDI of 80 %	MAP	SD
Riccardo	0.48	0.53	0.50	0.02
Kristian	0.99	1.0	1	0.005
Josh	0.80	0.81	0.80	0.003
Mikkel	0.49	0.50	0.51	0.004

Table 4 - Estimates of teachers' knowledge using the informed prior

Ruta Slivkaite
201805872@post.au.dk

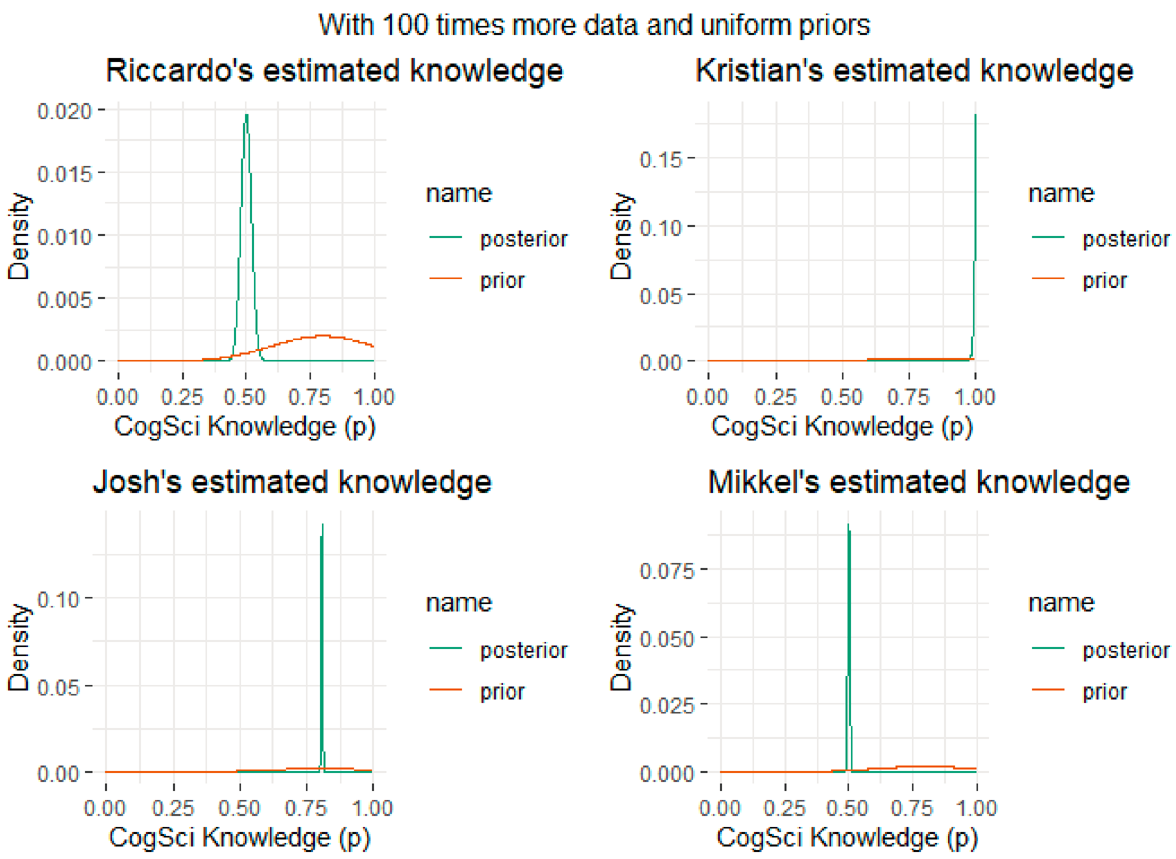


Figure 4 - Illustration of prior and posterior distribution of teachers' knowledge with uniform priors and 100 times more data

Ruta Slivkaite
201805872@post.au.dk

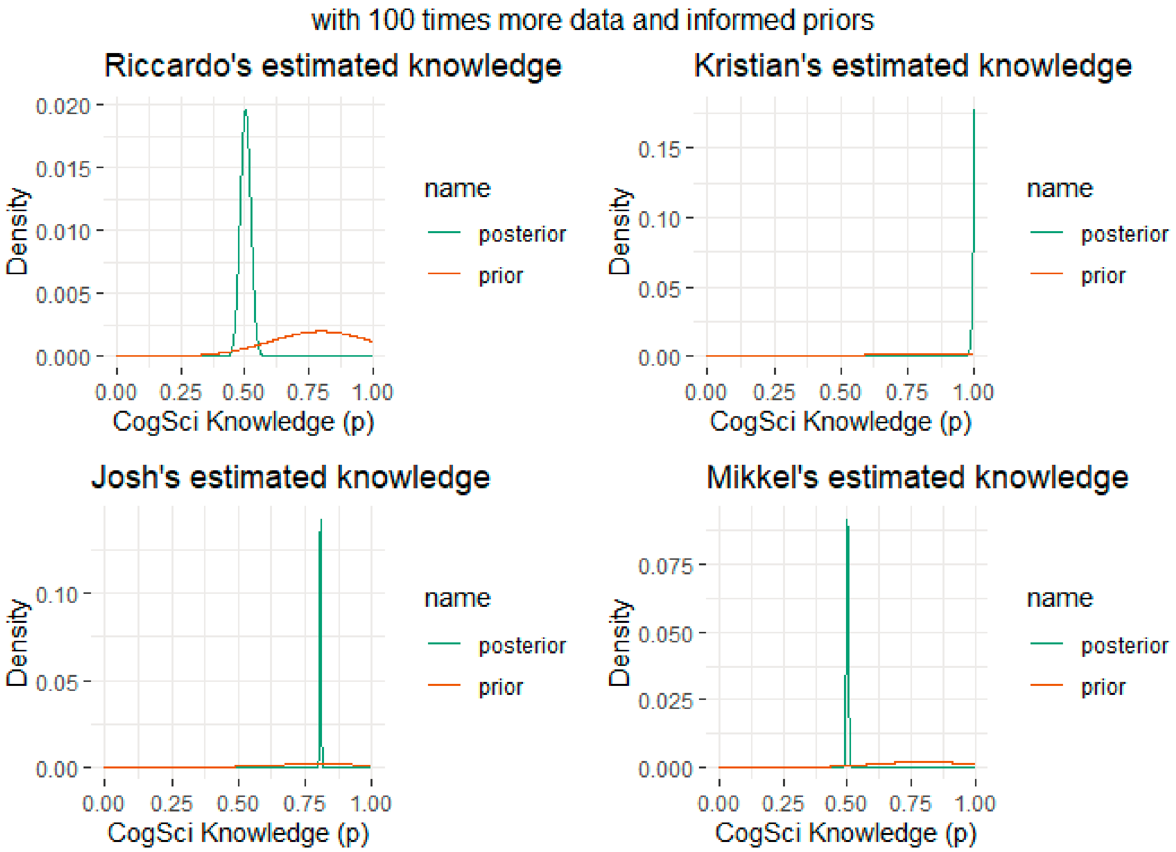


Figure 5 - Illustration of prior and posterior distribution of teachers' knowledge with informed priors and 100 times more data

5. Imagine you're a skeptic and think your teachers do not know anything about CogSci, given the content of their classes. How would you operationalize that belief?

If we were skeptical of our teachers' knowledge, we would need to adjust the prior accordingly. Assuming a normal distribution of teachers' knowledge, we would set a mean of the new prior to the lowest possible value: 0. Furthermore, we can reflect our skepticism in the standard deviation. The tighter it is, the more the model will be likely to pull the estimate down to zero, thus being skeptical. However, should we set it too strict, e.g., 0.001, we would probably run into divergence problems.

Ruta Slivkaite
201805872@post.au.dk

PART TWO

Last year we assessed the teachers' knowledge. This year we want to re-test them and assess whether our models are producing reliable predictions. Frequentist statistics use a machine-learning style to assess predictions (e.g. rmse on testing datasets). Bayesian stats makes things a bit more complicated.

1. Discussion of differences in the assessment of prediction performance in Bayesian vs. frequentist models

Assessment difference between **Bayesian** and **frequentist** models has its roots in the different ways the two interpret probabilities.

Frequentists consider probabilities as frequencies and they see data as a repeatable random sample (random variable) with a specific frequency/probability, which is defined as the relative frequency/probability of an event as the number of trials approaches infinity. In frequentists' view, events have a specific probability and the variation is in our sampling. Their model prediction performance is judged based on single-point estimates such as confidence intervals, hypothesis testing with p-values etc.

In contrast, **Bayesians** consider probabilities as degrees-of-belief, where a probability distribution is interpreted as quantifying our uncertainty about the world. In particular, this means that we can meaningfully talk about probability distributions of parameters, since even though the parameter is fixed, our knowledge of its true value may be limited. Bayesians assess model predictions appropriately measuring and accounting for the uncertainty in their models and parameters after looking at how a new piece of information has updated the prior and improved the certainty of the posterior.

2. Model prediction errors for the teachers given new data

Ruta Slivkaite
201805872@post.au.dk

Riccardo: In Riccardo's case, we can see from the histogram in figure 4 that the model almost never overpredicts Riccardo's knowledge. Instead it usually underpredicts his number of correct answers. We can see by the distribution that the model most often underpredicts by a value of 3.

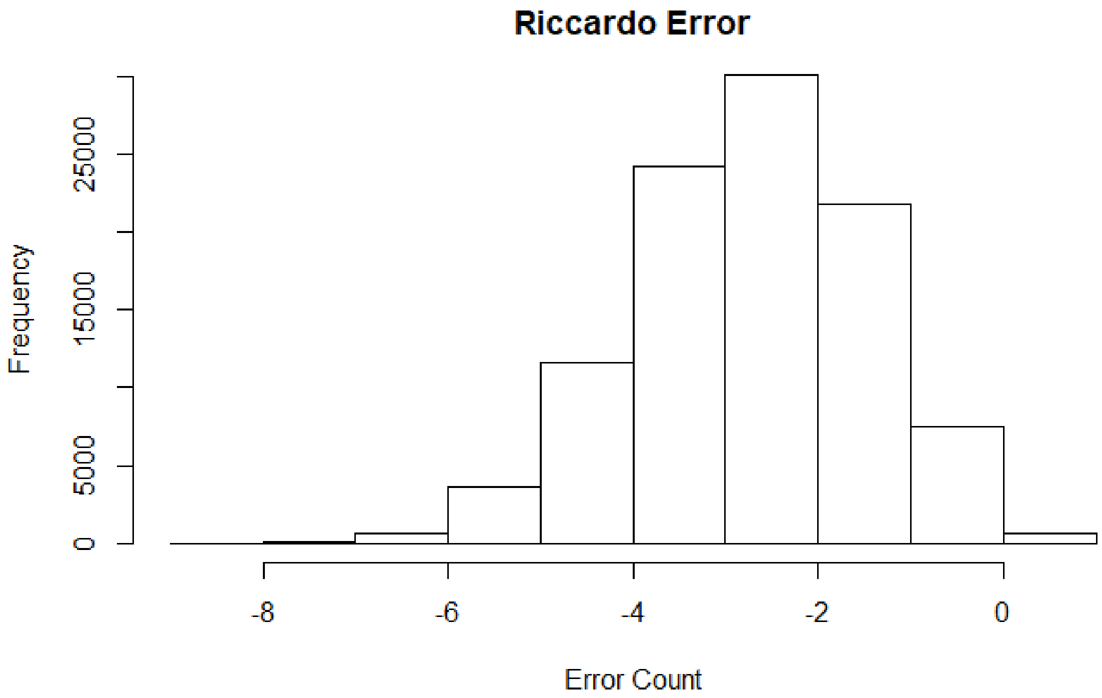


Figure 6 - The distribution of error from Riccardo's model given the observed data

Kristian: Looking at the histogram (figure 5) depicting the error of the model concerning Kristian's knowledge, we see that it both underpredicts and overpredicts according to the observed data. However, the model is most likely to overpredict, illustrated by the overweight of the distribution being above 0.

Ruta Slivkaite
201805872@post.au.dk

Figure 7 - The distribution of error from Kristian's model given the observed data

Josh: According to the histogram (figure 6) displaying the prediction error of the model for Josh, the model rarely overpredicts but often underpredicts his knowledge. Looking at the figure, we see that the model most often underpredicts by a value of around 10.

Figure 8 - The distribution of error from Josh's model given the observed data

Ruta Slivkaite
201805872@post.au.dk

Mikkel: As for Mikkel, figure 7 illustrates how the model is more likely to underpredict, although the model overpredicts as well. However, since Mikkel has answered 66 questions, and his error centers around -2, the underprediction is arguably less prominent than in the model's other cases.

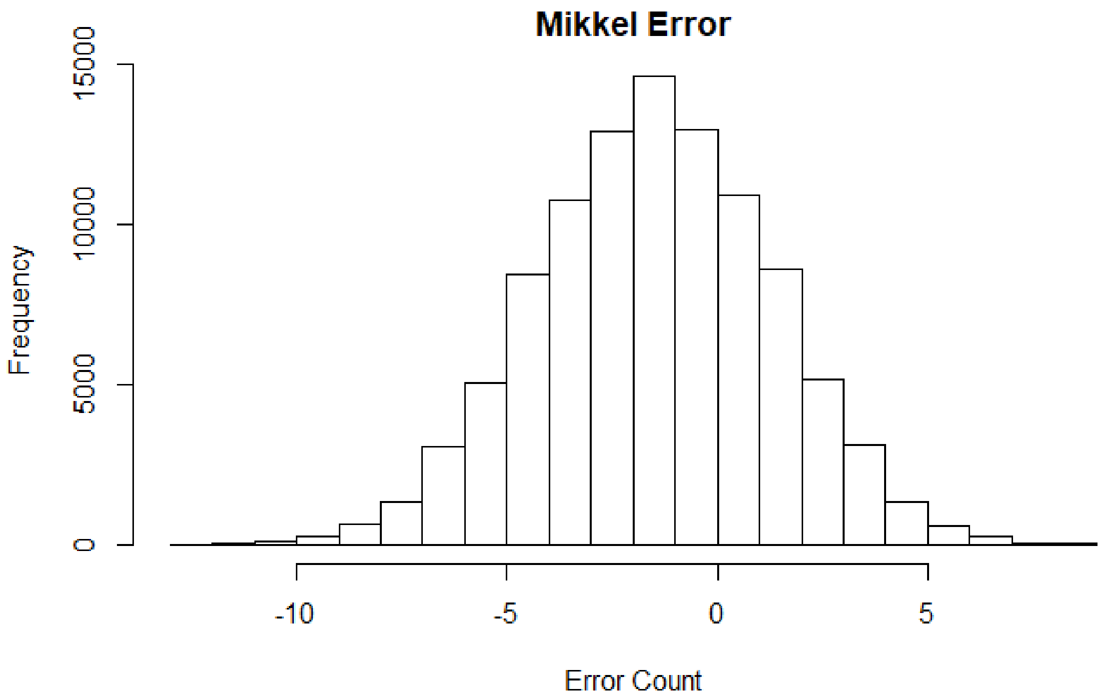


Figure 9 - The distribution of error from Mikkel's model given the observed data

Ruta Slivkaite
201805872@post.au.dk

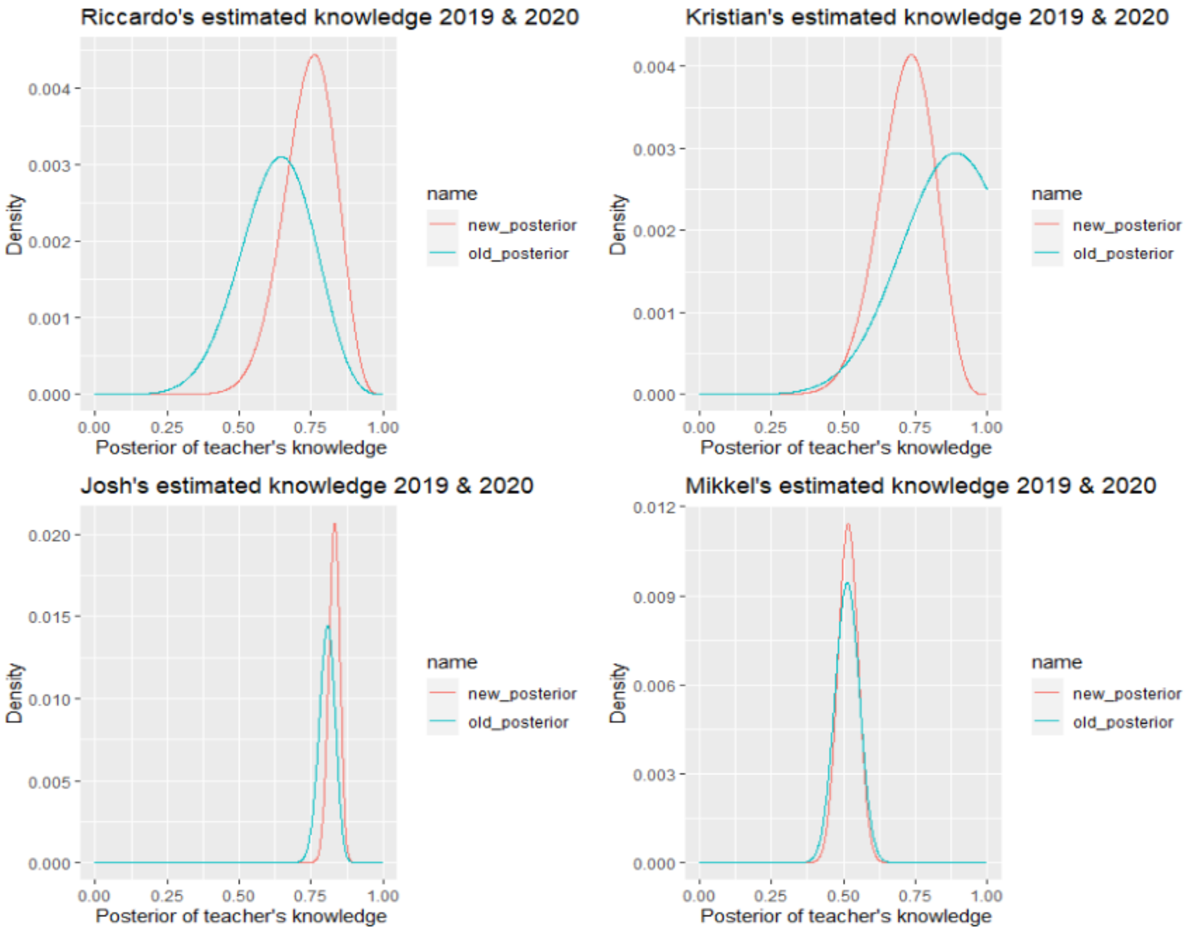


Figure 10 - Teachers' posterior given the old data compared to the updated posterior

Ruta Slivkaite
201805872@post.au.dk

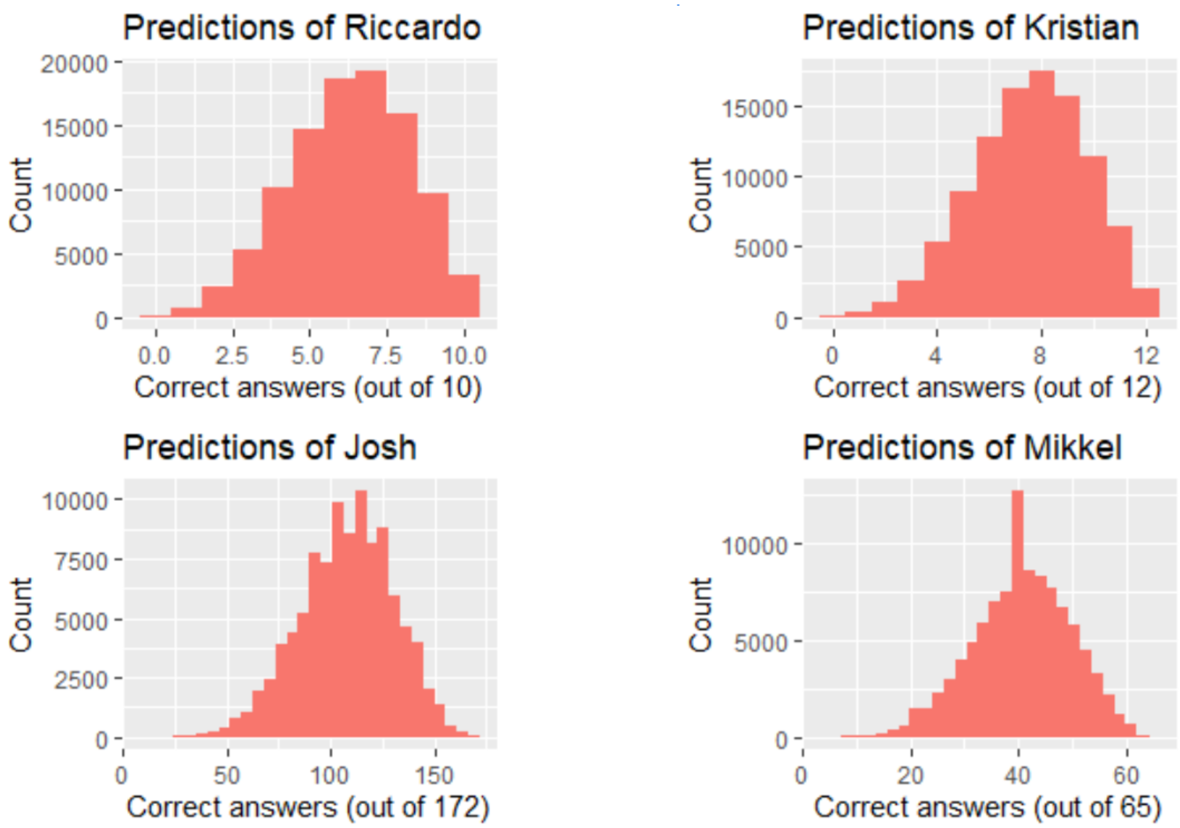


Figure 11 - Histograms displaying the predictions sampled from the posterior

Sources:

1. https://cxl.com/blog/bayesian-frequentist-ab-testing/
2. https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading20.pdf
3. https://www.probabilisticworld.com/frequentist-bayesian-approaches-inferential-statistics/
4. http://jakevdp.github.io/blog/2015/08/07/frequentism-and-bayesianism-5-model-selection/
5. https://www.probabilisticworld.com/intuitive-explanation-p-values/

Ruta Slivkaite
201805872@post.au.dk

Assignment 3

Causal Inference

GitHub Link: https://github.com/PeterThramkrongart/CompMod-3/blob/master/Assignment3_Peter-_Bella_Jakub_Bianka_Ruta.md

Introduction

Patients diagnosed with Schizophrenia are oftentimes found to possess altered control and distinction of self-other representations (e.g. hearing of voices or mind reading). Therefore this assignment was created with the aim of investigating this matter. In particular, we have examined altercentric intrusions (inability to inhibit social information) to see whether it can truly be associated with Schizophrenia and these aforementioned psychotic symptoms. In addition, we have also checked whether apathy (severe lack of motivation in taking care of oneself, e.g. getting out of bed or showering to show up for work) could be in relation with altercentric intrusions. In order to test our research questions we built multiple regression models. The data used in this assignment was simulated data. This is because the intention of this assignment is to explore some of the biases that can be introduced and the effects that can be unmasked when using multivariate modelling.

Hypotheses

In this paper we aim to investigate two distinct hypotheses:

H1: Schizophrenia is related to altercentric intrusions

H2: Altercentric intrusion is related to Schizophrenic symptoms (i.e. delusions of voice hearing & mind reading)

Ruta Slivkaite
201805872@post.au.dk

PART ONE

1.1. Altercentric intrusion in schizophrenia and controls

Model and results

In order to check the hypothesis that altercentric intrusion is higher in schizophrenia patients compared to controls, we built a univariate regression model with altercentric intrusion as the outcome and diagnosis as a predictor.

Model 1: Altercentric Intrusion diagnosis model

Altercentric Intrusion ~ 0 + Diagnosis

Since we are utilizing Bayesian modeling, in addition to designing the model, we assign priors for all of the parameters in the model. These priors define the initial information state of the model, before seeing the results. The prior for sigma representing the expected average error of our model is deduced from a standard deviation of altercentic intrusion. The beta prior was constructed to accommodate the entire distribution of altercentric intrusion, but still having a regularization function to prevent from overfitting, (see the Figure 14 for the prior check in the appendix).

Priors:

β ~ Normal (4, 1)

σ ~ Normal (1, 2)

Ruta Slivkaite
201805872@post.au.dk

The results of the model indicate a credible difference in altercentric intrusion in the two groups, supporting our hypothesis (b = 0.36, CIs = 0.16, 0.56, ER = 665.67). The difference in spread of data is depicted in Figure 1.

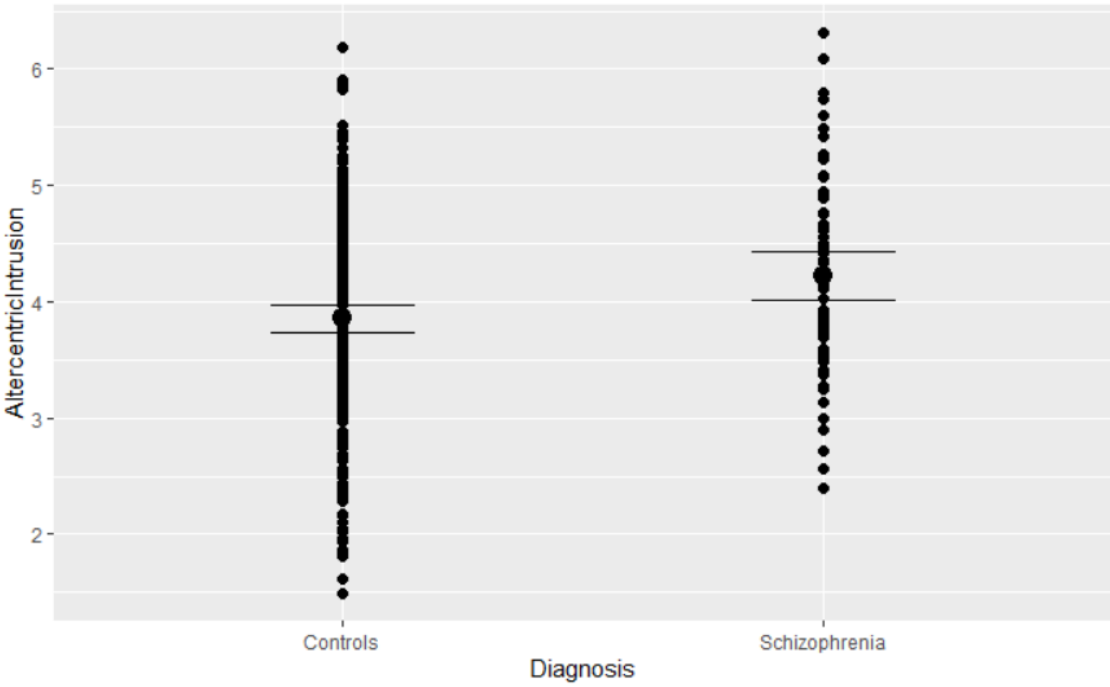


Figure 1: Illustration of altercentric intrusion in both Schizophrenia and Controls

Controls showed an on average altercentic instrusion effect of 3.86 (CIs = 3.75, 3.98) and schizophrenia of 4.22 (CIs = 4.02, 4.43). The model had no divergences, a Rhat of 1, and an effective sample size above 2000 for both Bulk and Tail.

The results of the analysis indicates that it is credible that altercentric intrusion is more prevalent among people with schizophrenia than among controls. This is evidence in favor of our first hypothesis, H1. Looking at the posterior predictive check (see Figure 15 in the appendix), we can see that your model given the predictor is fairly accurate at predicting new data, hence we can conclude that model makes sense. The update check plot (see in the Figure 16 in the appendix) shows the improvement from posterior, initial knowledge, to the posterior after fitting the model to data. As the posterior shrinks, it becomes more certain and learns more from the data.

1.2. Altercentric intrusion relation to separate symptoms in schizophrenia

Ruta Slivkaite
201805872@post.au.dk

Models and the results

In order to figure out which symptoms are related to altercentric intrusion in schizophrenia patients, we built univariate regression models and assigned priors for all of the parameters in the models. As a part of the preprocessing, we scaled the data so that the symptom severities would be comparable. Furthermore, we only used the data from the diagnosed individuals due to the nature of the question.

In the analysis, we model altercentric intrusions as the outcome and the separate symptoms, namely, voice hearing, mind reading, and apathy, as predictors. Moreover, to unmask any shared variance, we include two multivariate models.

Due to the fact that all variables were generated on the same scale, we are able to use identical priors across all models. The prior was designed following both conceptual knowledge and trial-and-error learning with the aim of model's regularization. The intercept prior was set at zero as we scaled the data. In terms of sigma, we deduced the value from a distribution of altercentric intrusion. Lastly, the beta prior was centered at zero due.

We have run the prior and posterior predictive checks, also, updating checks with hypothesis testing for all models, for obtained plots see the appendix Figures 17-29. For all of the models the posterior is more certain than the priors, therefore we can assume that the models have learned something from the data.

Priors:

α ~ Normal (0, 1)

β ~ Normal (0, .3)

σ ~ Normal (1, 2)

Model 2: Voice Hearing model

Altercentric Intrusion ~ 1 + Voice Hearing

The outcome of model 2 indicates a negligable positive effect of voice hearing on altercentric intrusion in suport of our hypothesis (b = 0.07, CIs = -0.15, 0.3, ER = 2.28). The model had no

Ruta Slivkaite
201805872@post.au.dk

divergences, a Rhat of 1 and Effective sample sizes above 2000 for both Bulk and Tail. The regression line is depicted in figure 2.

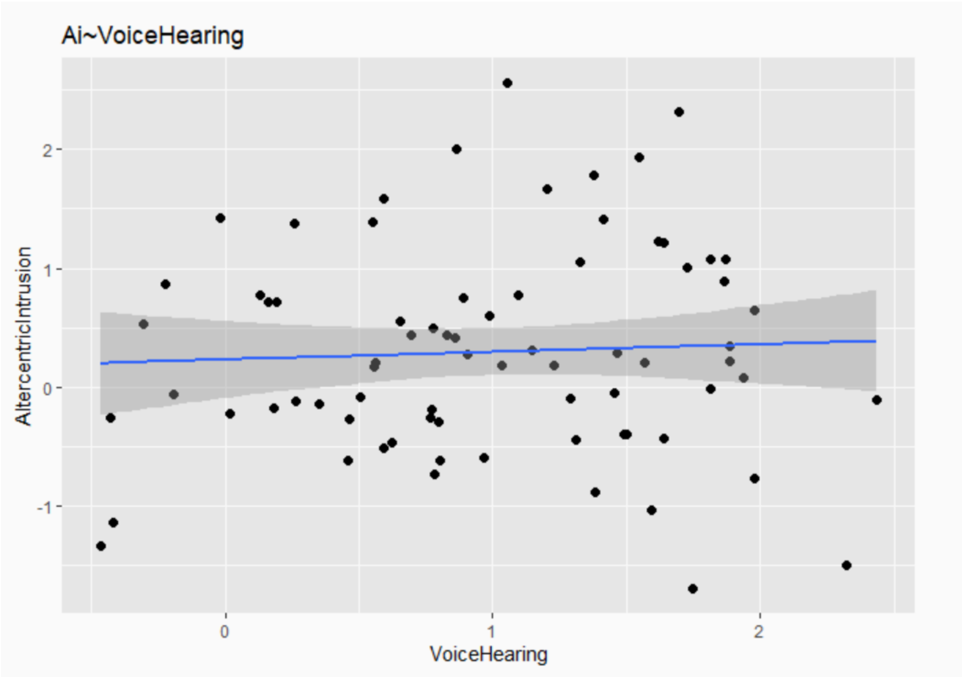


Figure 2: Plotting a regression line against the data

Model 3: Mind Reading model

Altercentric Intrusion ~ 1 + Mind Reading

Model 3 indicates a negligable positive effect of mind reading on altercentric intrusion, suporting our hypothesis (b = 0.07, CIs = -0.11, 0.25, ER = 2.87). The model had no divergences, a Rhat of 1 and Effective sample sizes above 2000 for both Bulk and Tail. The regression line is depicted in figure 3.

Ruta Slivkaite
201805872@post.au.dk

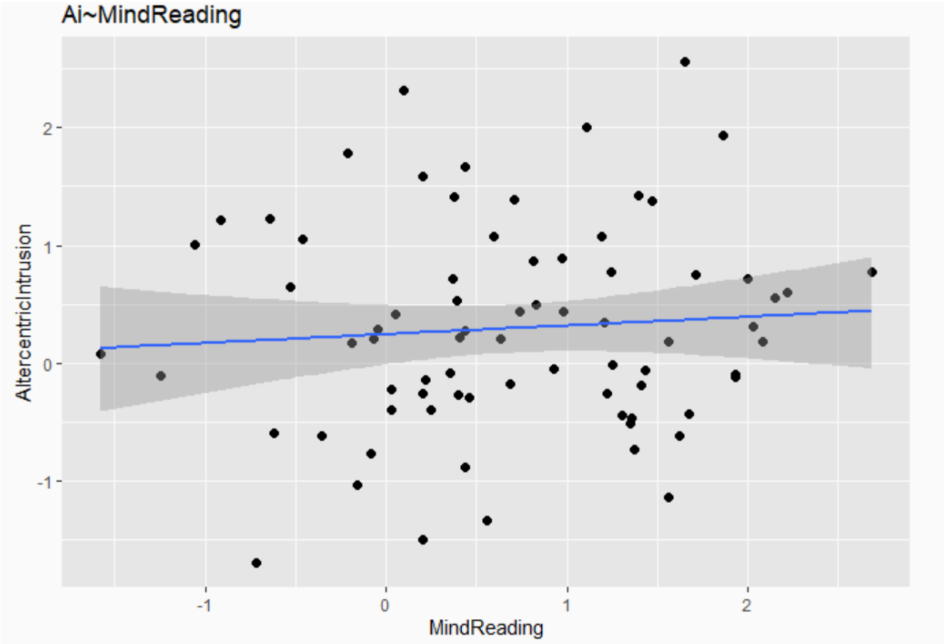


Figure 3: Plotting a regression line against the data

Model 4: Apathy model

Altercentric Intrusion ~ 1 + Apathy

Model 4 indicates a credible negative effect of apathy on altercentric intrusion, which seems to be against our hypothesis (b = -0.23, CIs = -0.50, 0.04, ER = 3.66), which predicts no effect of apathy on AI. The model had no divergences, a Rhat of 1 and Effective sample sizes above 2000 for both Bulk and Tail. The regression line is depicted in figure 4, showing a negative correlation between apathy and AI.

Ruta Slivkaite
201805872@post.au.dk

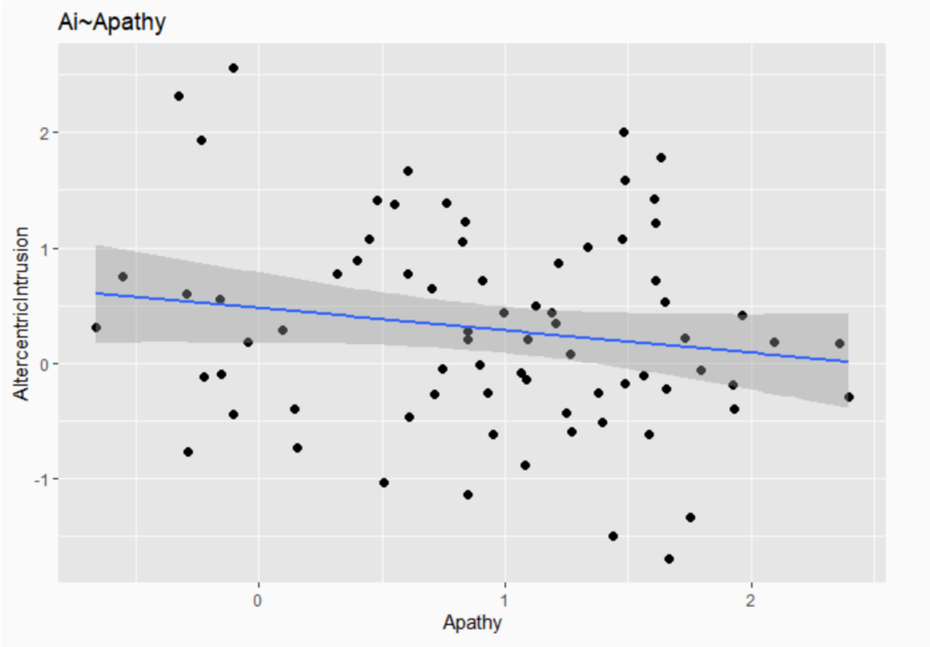


Figure 4: Plotting a regression line against the data

Model 5: Mind reading and Voice hearing model

Altercentric Intrusion ~ 1 + Voice Hearing + Mind Reading

Model 5 is a multivariate model, which examines how much of the variance either voice hearing or mind reading can explain on its own, when the other symptom is taken into account. The outcome indicates a slight effect of voice hearing on altercentric intrusion (b = O.11, CIs = -0.19, 0.39, ER = 3.42), as well as a slighter effect of mind reading on AI (b = 0.10, CIs =-0.13, 0.32, ER = 4.46). See figure 5. The model had no divergences, a Rhat of 1 and Effective sample sizes above 2000 for both Bulk and Tail. Granted, while these results are pointing in the direction of our hypothesis, they are neither particularly satisfying nor convincing.

Ruta Slivkaite
201805872@post.au.dk

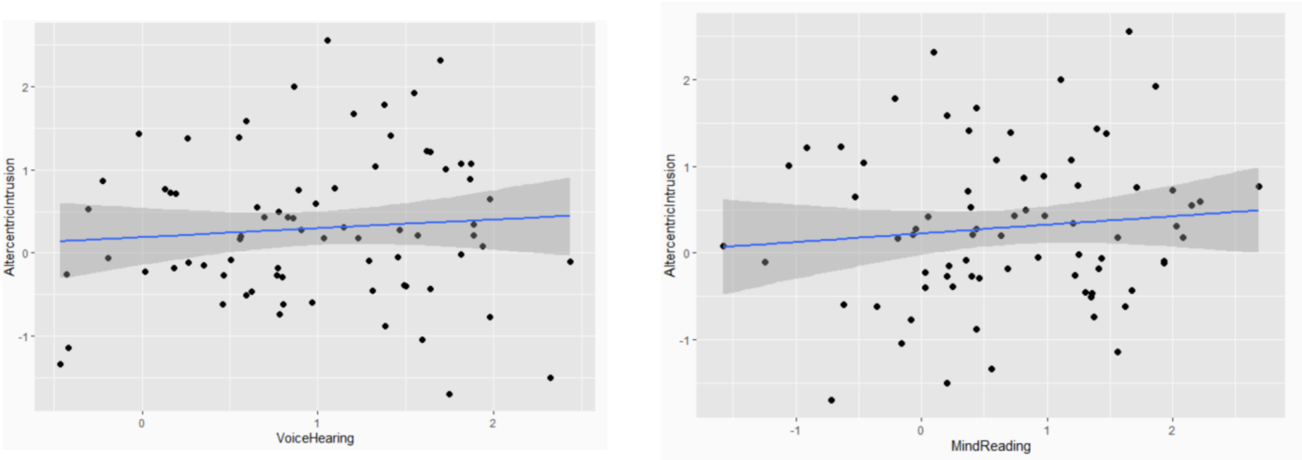


Figure 5: Plotting a regression line against the data

In order to see how much each of the predictors is contributing to explaining the altercentric intrusion when they are modeled together, we built a multivariate regression model with all of the symptoms as predictors and altercentric intrusion as an outcome. The data still consists of solely positively diagnosed individuals.

Model 6: Mind reading, Voice hearing and Apathy model

Altercentric Intrusion ~ 1 + Voice Hearing + Mind Reading + Apathy

The final model in this section, model 6, indicates a barely credible positive effect of voice hearing and mind reading, which supports our hypotheses, and a slightly more credible negative effect of apathy, which rejects our hypothesis once more: mind reading (b = 0.04, CIs = -0.17, 0.24, ER = 1.58), voice hearing (b = 0.05, CIs = -0.21, 0.3, ER = 1.77), apathy (b = -0.17, CIs = -0.42, 0.07, ER = 7.25). The model had no divergences, a Rhat of 1 and Effective sample sizes above 2000 for both Bulk and Tail. See figure 6 for regression lines from individual predictors.

Ruta Slivkaite
201805872@post.au.dk

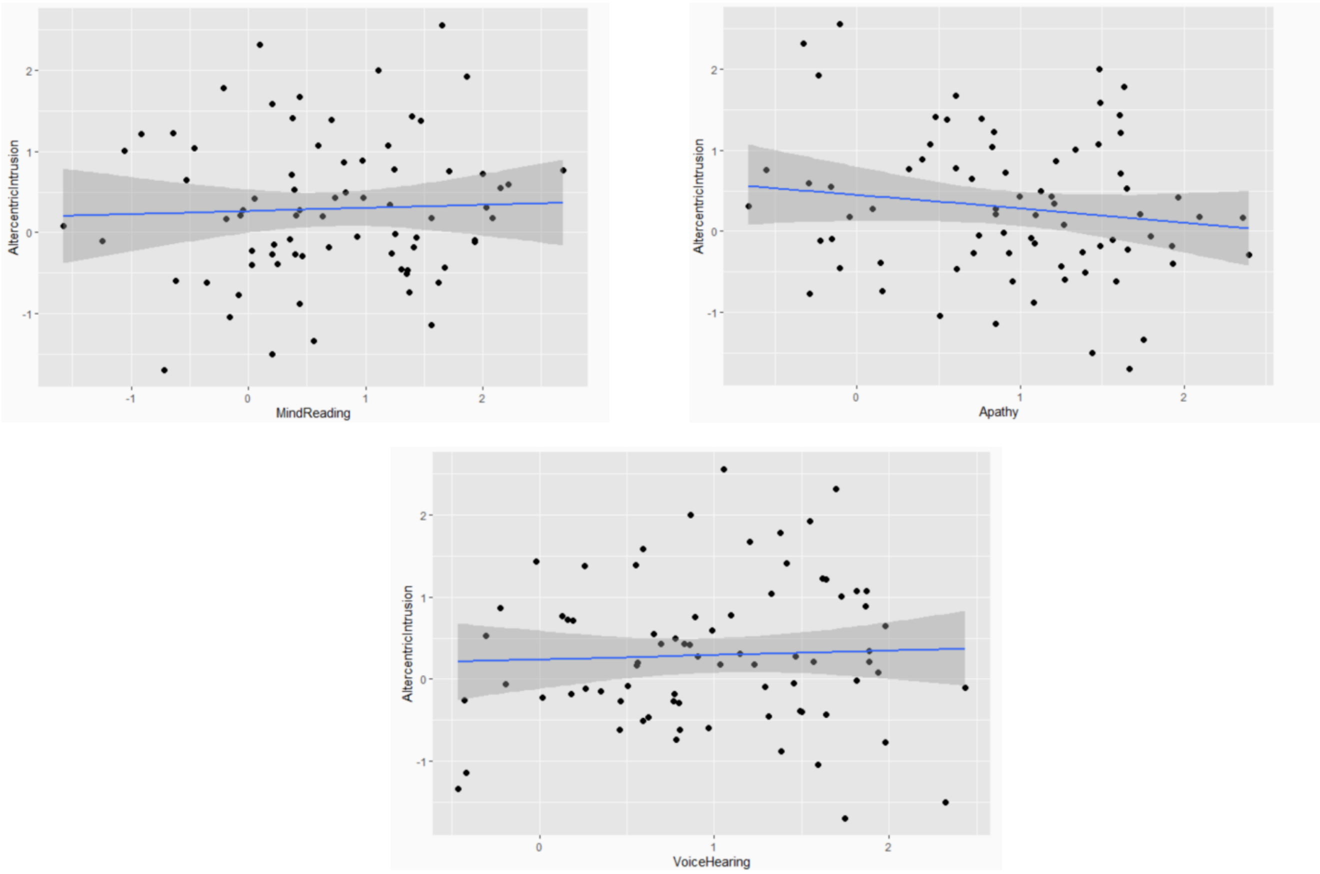


Figure 6 :Plotting a regression line against the data

We do a model comparison of our best models- which we argue to be model 6 and model 3 based on the significance of effects- in order to check, which model performs better at predicting altercentric intrusion. So as to achieve this we utilize a leave-one-out cross-validation. The results are shown in table 1 below. The model comparison indicates that predicting AI from apathy is the optimal course of action, if we want to minimize estimated out of sample error (stocking weight of 1). We see that adding other symptoms to the apathy model does not improve the generalizability of the model (stocking weight of 0).

Ruta Slivkaite
201805872@post.au.dk

Model	elpd difference	SE difference	Weight
M4, Apathy	0	0	1
M3, Mind Reading	-1.1	1.3	0
M2, Voice Hearing	-1.3	1.4	0
M6, VH + MR + A	-1.6	0.4	0
M5, VH + MR	-1.6	1.2	0

Table 1: Leave-one-out cross validation (schizophrenics only)

The immediate interpretation of these results is that in patients with schizophrenia, apathy seems to predict altercentric intrusion the best out of the three symptoms analyzed.

Discussion

According to our analysis and leave-one-out comparison, apathy is clearly the best predictor of altercentric intrusion out of the three symptoms, and the model does not improve significantly by adding the two other symptoms. The results of the given models thus go against our theoretical knowledge and, therefore, our hypotheses that mind reading and voice hearing are related to altercentric intrusion. Instead, we see that apathy is a more credible predictor, which is against our hypothesis. In the light of this evidence, we need to rethink our theoretical assumptions.

However, instead of readily rejecting our hypothesis, which is grounded in theory, we might attempt to analyze the data more thoroughly first, in order to exclude the possibility of statistical biases and confounds. In the next section we attempt to unravel biases in the analysis.

PART TWO

Ruta Slivkaite
201805872@post.au.dk

2.1. Investigating Model Biases

In order to investigate the data and the implicit inferences we make, when designing the analysis, we illustrate our best conjecture for the correlations between the symptoms and the diagnosis of schizophrenia in a directed acyclic graph (DAG) (see figure 7).

According to the DAG, which is congruent with our hypothesis, only voice hearing and mind reading are causally associated with altercentric intrusion, while all three symptoms together elicit a diagnosis.

In the small world the DAG presents, where the schizophrenia diagnosis is given by an additive combination of these three symptoms, people with the diagnosis are likely to either have medium to high values in voice hearing and mind reading, or high values in apathy. If apathy and voice hearing are not significantly correlated, it is improbable to score high in both categories, but probable to show one of two profiles. You could even count three profiles, if you wanted to dissociate VH and MR, but it is not necessary to explain the significant negative correlation between apathy and altercentric intrusion.

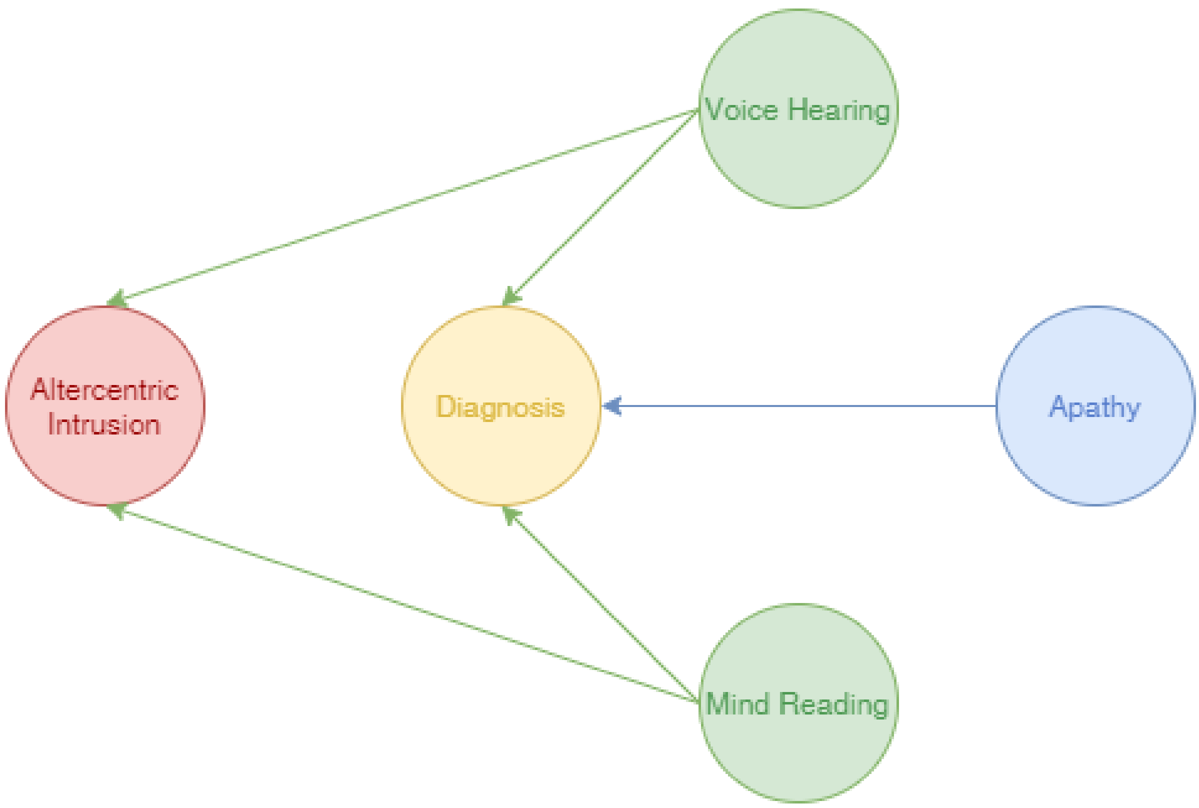


Figure 7: DAG illustrating our hypothesized causal relationships

Ruta Slivkaite
201805872@post.au.dk

Therefore the reason for apathy being a better predictor of altercentric intrusion than mind reading or voice hearing, despite the fact that altercentric intrusion is not causally related to apathy, is that the analysis deals with a diagnosed group of people, which are thus selected non-randomly from the whole dataset. If we examine how the diagnosis is administered, by looking at the data simulation, we see that people are assigned the diagnosis by being in the top 25 % according to the sum severity of three symptoms: apathy, mind reading, and voice hearing.

In a scenario where the three symptoms are independent (or, perhaps, weakly correlated), and where we only look at people with schizophrenia, apathy actually *should* be highly negatively correlated with altercentric intrusion for the very reason that they are *not* causally related in the DAG. Individuals who only experience weak voice hearing and mind reading, but are still deemed Schizophrenic, according to our simulator, *must* have relatively high scores in apathy. There are only those three symptoms. Since we only analyze diagnosed individuals, many of these individuals with high apathy scores will in contrast have weak scores for AI, thus creating the spurious correlation.

In our case the predictor variables are not independent, which makes inferring about biases more complicated. All predictors are correlated with each other, but equally and only weakly. Therefore we believe the bias is still present, and we can thus expect to see a negative correlation between apathy and altercentric intrusion within the Schizophrenia diagnostic group. We believe this bias originates from diagnosis being an addition of multiple variables. This means that uncorrelated or positively correlated variables like voice hearing and apathy become negatively correlated within a patient group, thus making anything dependent on e.g. voice hearing, spuriously correlated with apathy as well.

If we are interested in examining the symptomatic correlations to altercentric intrusion, we suggest using both parts of the dataset: diagnosed and undiagnosed. Another approach could be to include more symptoms, so that the diagnosis isn't only caused by either the altercentric intrusion complex or apathy.

2.2. Rethinking the Analysis

When redesigning our analysis, we want to use some of the ideas and inferences from the section above. Since we have argued that selecting only schizophrenic patients creates a spurious correlation between apathy and altercentric intrusion, it seems intuitive to retry some of the original models using the whole dataset. If we do not see the same negative relationship between apathy and altercentric intrusion in the whole data set as in diagnosed individuals, we can be pretty certain that our results are influenced by some sort of bias.

Ruta Slivkaite
201805872@post.au.dk

The results of the models 2-6 as described in pseudocode under 1.2., when rerun on the entire data set of schizophrenic patients and controls alike, are shown in table 2 below. In order to attain these results, we run uni and multivariate regressions, reusing the priors from 1.2., with altercentric intrusion as the dependent variable. Predictor variables are shown in *Table 2*.

	b	**CIs**	**ER**
Model 2			
Voice Hearing	0.19	0.08, 0.30	Inf
Model 3			
Mind Reading	0.19	0.07, 0.30	1999
Model 4			
Apathy	0.08	-0.03, 0.20	12.75
Model 5			
Voice Hearing	0.16	0.06, 0.27	570.43
Mind Reading	0.16	0.06, 0.27	799
Model 6			
Voice Hearing	0.16	0.05, 0.27	306. 69
Mind Reading	0.16	0.05, 0.27	799
Apathy	0.01	-0.09, 0.12	1.57

Table 2: output of six separate bayesian analyses

Just by surveying the outcome of the analyses, it is evident that the large effect of apathy on altercentric intrusion has virtually disappeared, when we add controls to the analysis. On the other

Ruta Slivkaite
201805872@post.au.dk

hand the credibility and evidence ratio as well as the estimates of voice hearing and mind reading being related to AI has increased.

In order to conclude which model performs the best, when predicting altercentric intrusion in individuals, we compare models using the LOO method, *see Table 3* below.

Model	elpd difference	SE difference	Weight
M5, VH + MR	0	0	0.806
M6, VH + MR + A	-1	0.3	0
M2, Voice Hearing	-3.5	2.9	0.080
M3, Mind Reading	-3.6	3	0.114
M4, Apathy	-8.2	4.3	0

Table 3: Leave-one-out cross validation (all data)

The results of the analysis are indicated in *Table 3* above. According to our LOO cross-validation analysis, the model predicting altercentric intrusion from the combinated symptom severity of mind reading and voice hearing is superior in performance with a stocking weight of 0.806. Model 5 (VH + HR) is thus the model which best minimizes error in predicting AI. See figures 11 and 12 for visualization of relation of the two predictors to altercentric intrusion according to model 5.

When comparing the univariate regressions (Model 2 and 3) to the multivariate regression (Model 5), the results seem to indicate a slight overlap in the variance that mind reading and voice hearing explains. However, the symptoms are still able to explain separate parts of the variance, lending credibility to our assumption in the DAG, that both symptoms influence AI separately.

The results from the rerun of the analysis on the entire dataset is in favour of our original hypothesis, namely that altercentric intrusion is related to a combination of factors, particularly mind reading and voice hearing. The new analysis paints an entirely different picture of our dataset than our first attempt. It illustrates the importance of knowing your data, analysis, and assumptions, in order to weed out conceptual confounds.

Ruta Slivkaite
201805872@post.au.dk

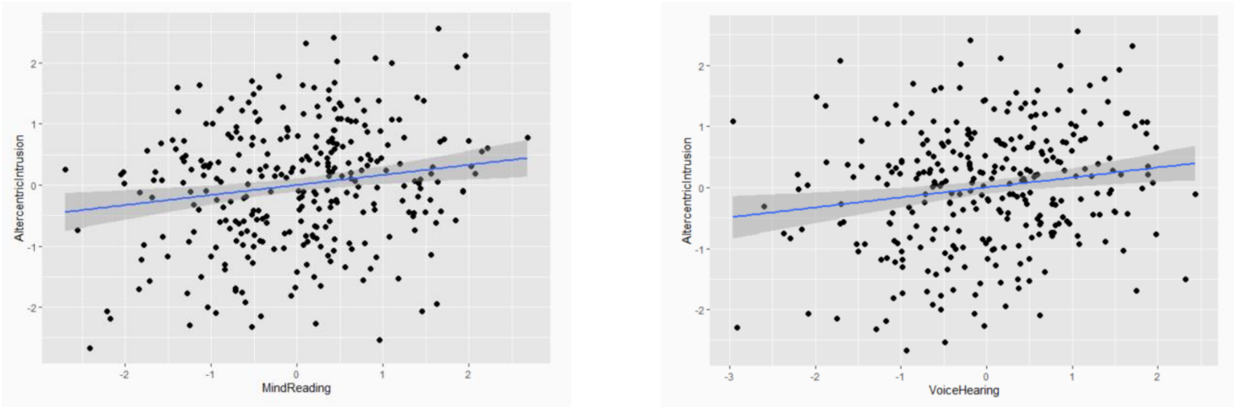


Figure 8: Plotting a linear regression line against the data, Model 5

Figure 9: Visualization of our preferred model, Model 5

PART THREE

3.1 Simulation of data

When exploring how the data was generated, we can see that diagnosis is defined as a sum of all three symptoms with a threshold of 75 %. In other words, only the 25 % of all “people” with the highest sum of symptoms are classified as schizophrenic. Moreover, a weak correlation of 0.2 was set as symptoms usually co-occur with one another. In the simulation, altercentric intrusion was created to be 20 % correlated with both mind mind reading and voice hearing. This enables us to test whether the inclusion or exclusion of controls in our data is the cause of spurious correlations.

Ruta Slivkaite
201805872@post.au.dk

When we compare the correlations between variables of schizophrenia, we find that the inclusion of the control group leads to relations between variables similar to those defined in the variance-covariance matrix for generating the simulated data (*see Figure 12*).

On the contrary, the exclusion of the control group logically creates negative relations between variables (*see Figure 10*) because if one variable is high enough and causes a participant to be classified as schizophrenic, then the other symptoms are more likely to be lower. These results are in line with the findings in part 1.2 (*see Figures 14 and 16)* and, also, lead to apathy being the best predictor of AI, which is against our theoretical assumptions. Therefore, it is a great lesson of how knowing and grounding our assumptions in theory is an important part of making causal inferences without introducing statistical confounds and biases.

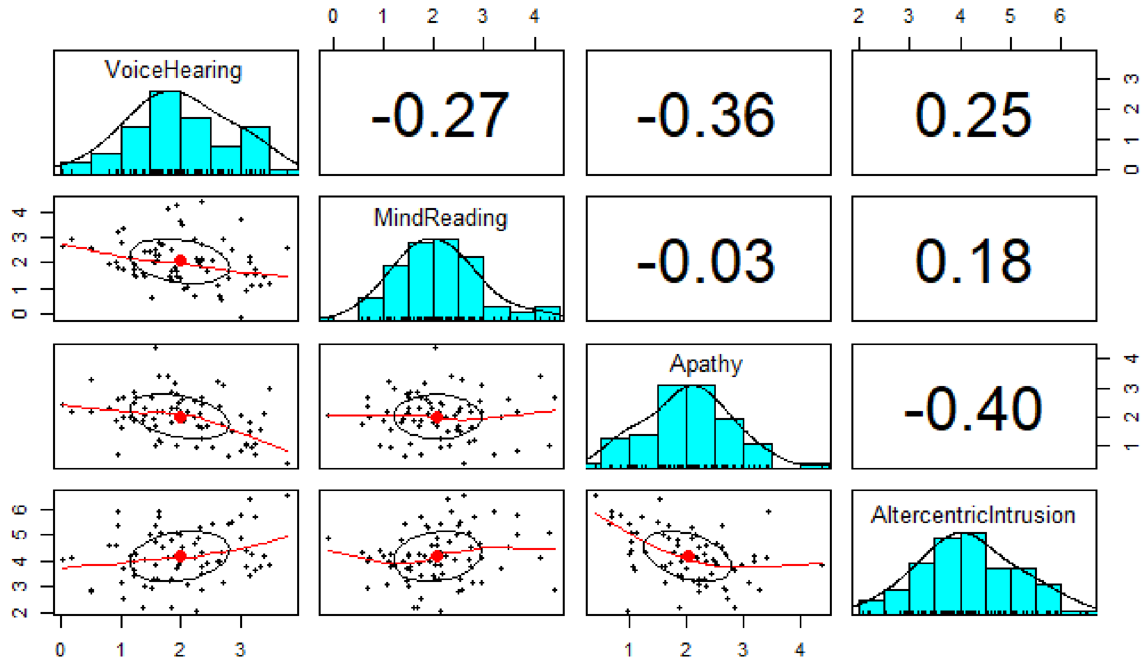


Figure 10: variance-covariance matrix of relations between variables in a simulated schizophrenia data

Ruta Slivkaite
201805872@post.au.dk

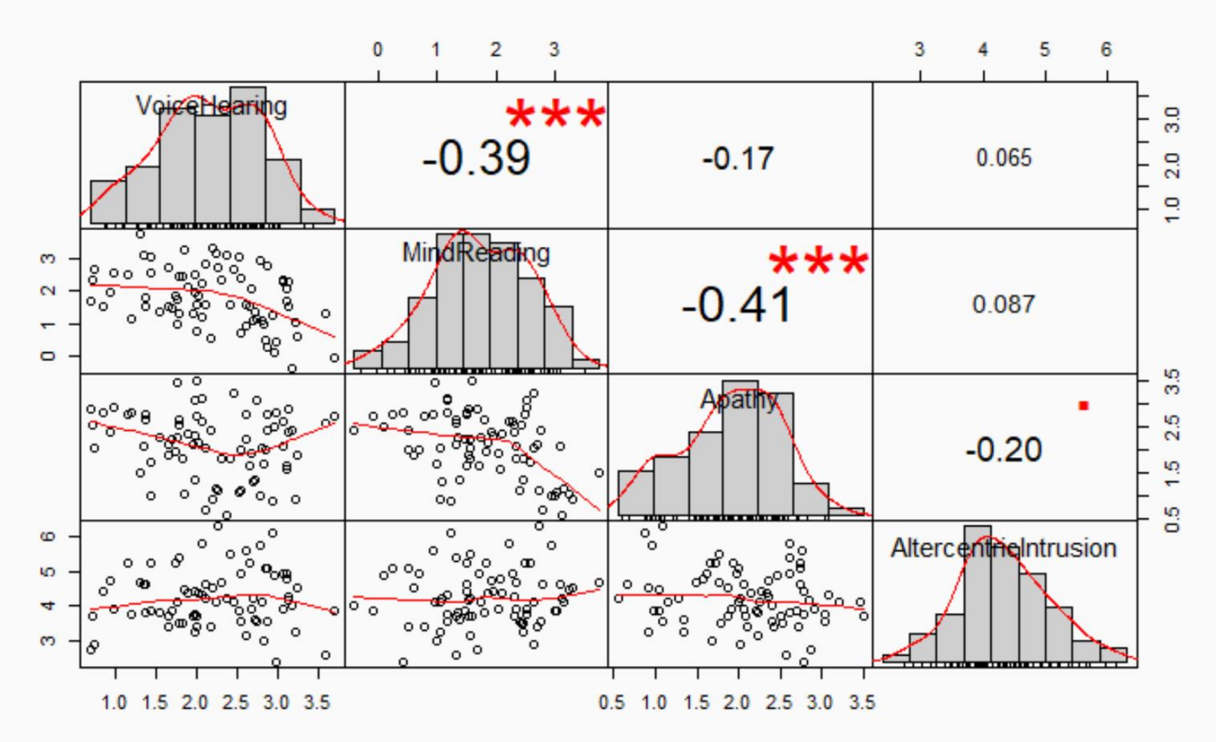


Figure 11: variance-covariance matrix of relations between variables in tested schizophrenia data

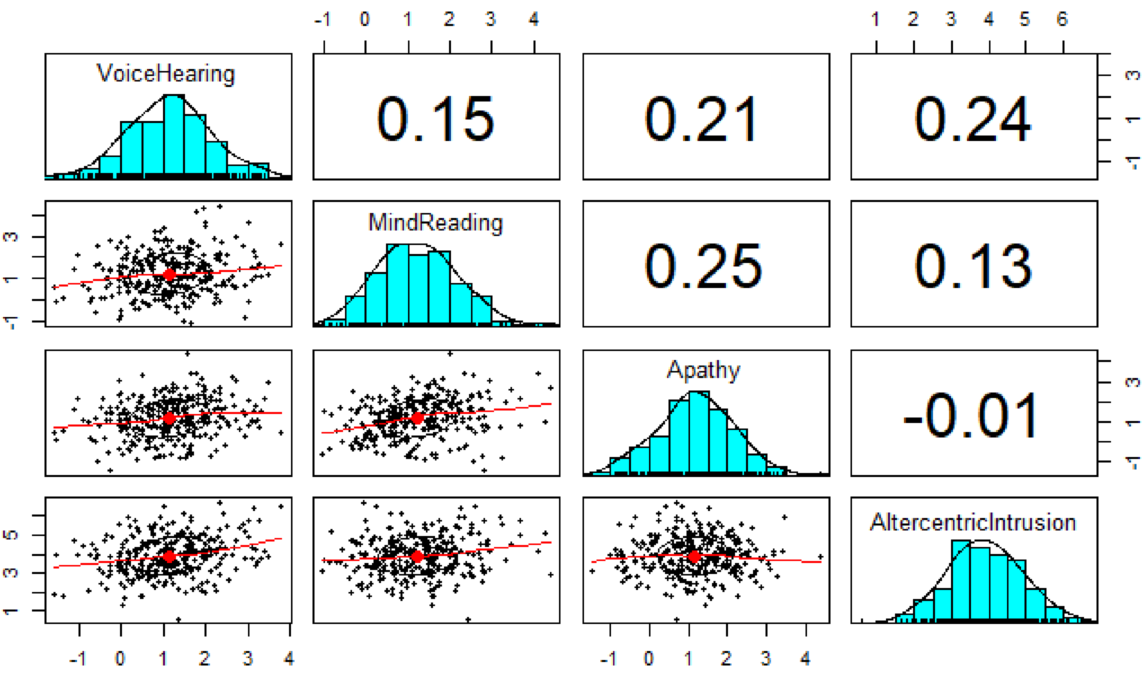


Figure 12: variance-covariance matrix of relations between variables in simulated controls data

Ruta Slivkaite
201805872@post.au.dk

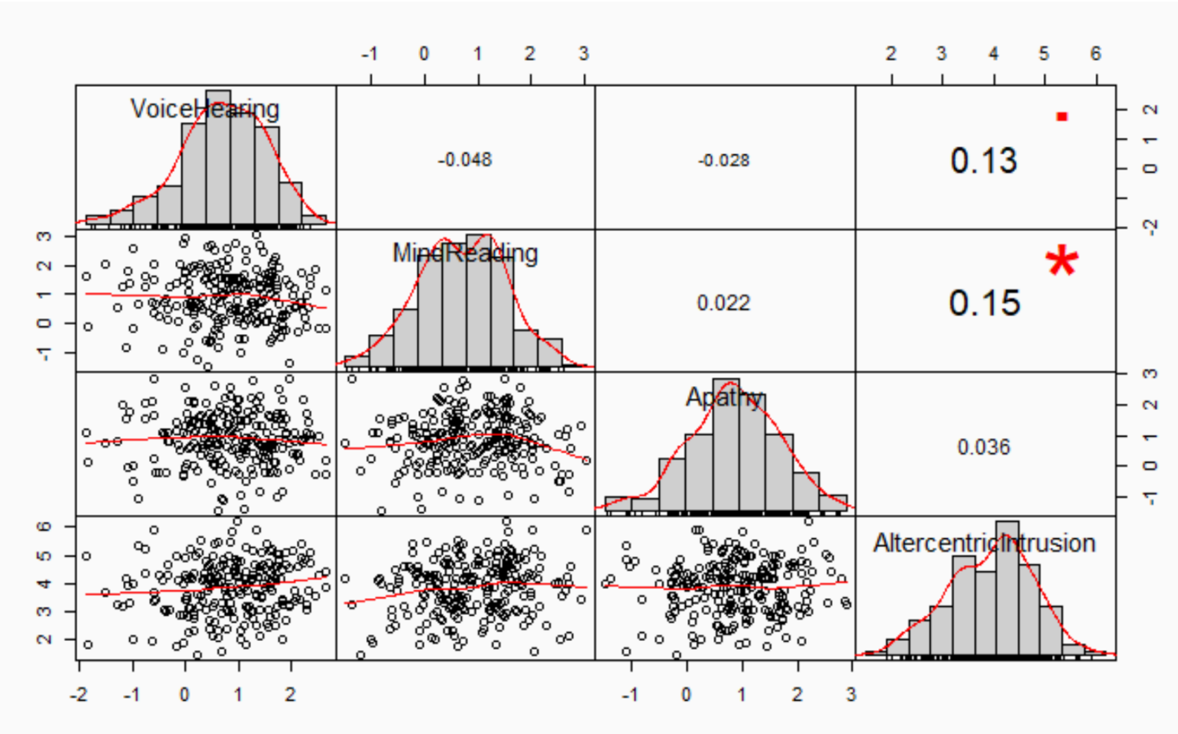


Figure 13: variance-covariance matrix of relations between variables in tested controls data

Ruta Slivkaite
201805872@post.au.dk

Appendix

Part One

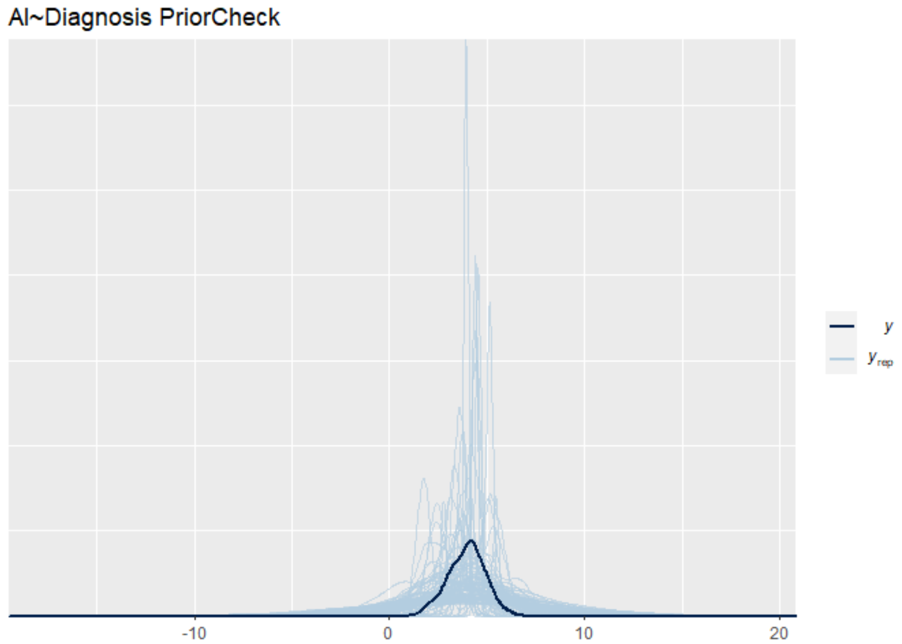


Figure 14: prior predictive check

Ruta Slivkaite
201805872@post.au.dk

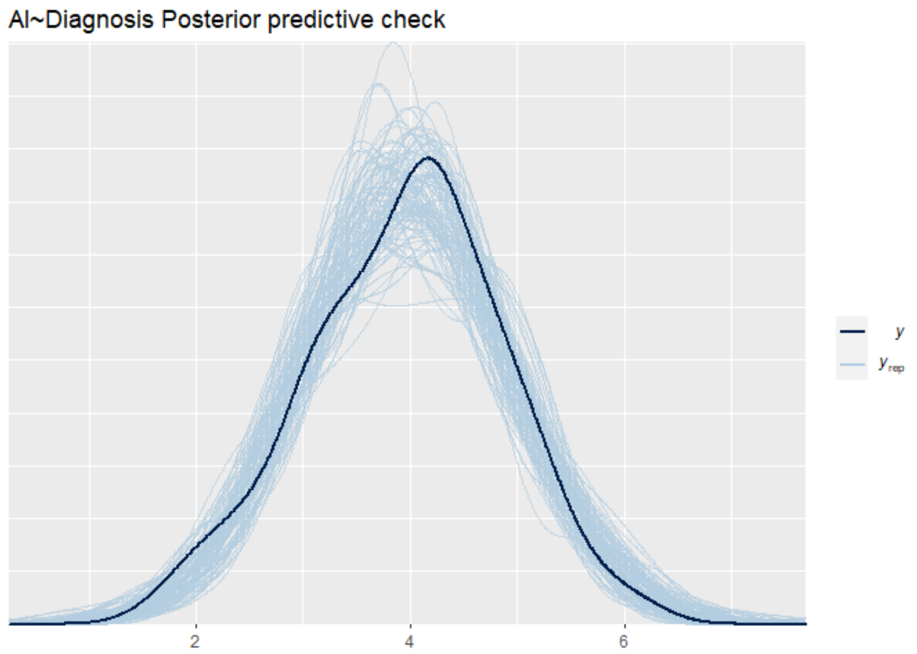


Figure 15: posterior predictive check

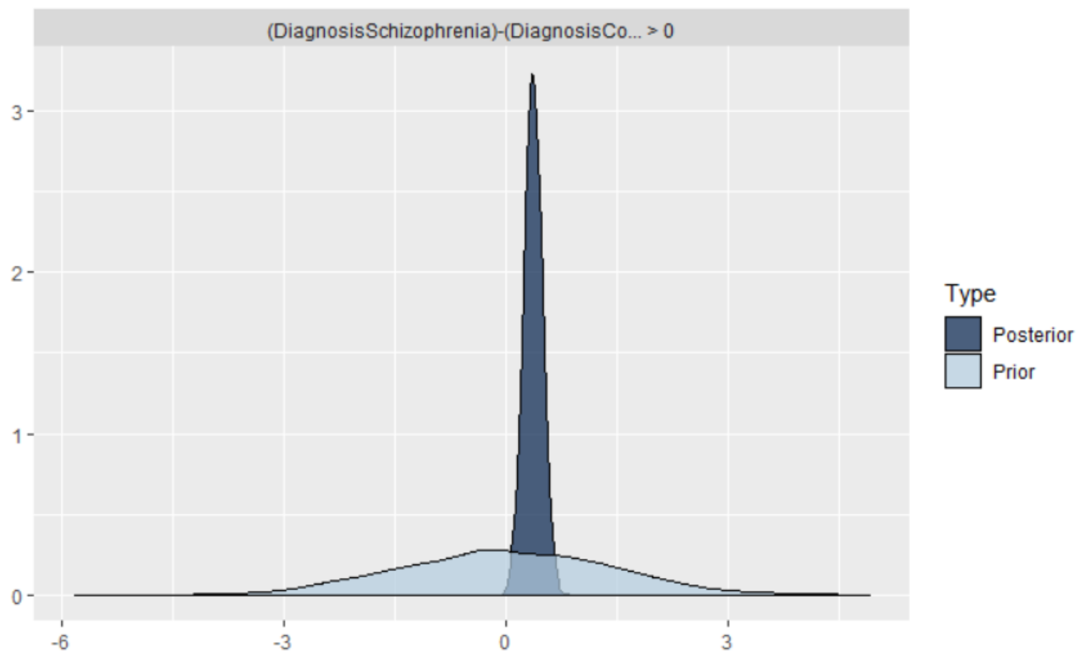


Figure 16: update check plot + hypothesis testing

Ruta Slivkaite
201805872@post.au.dk

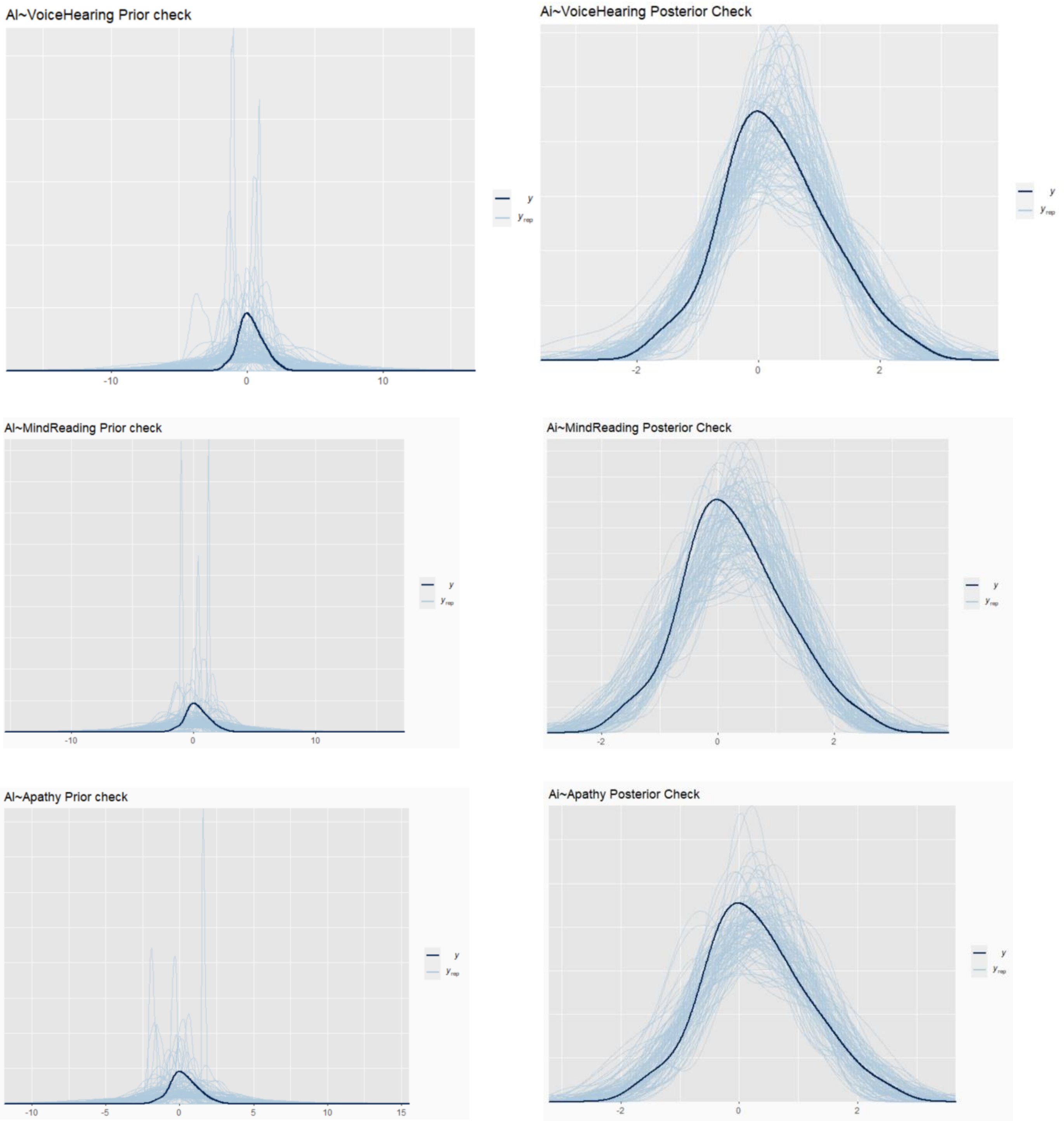


Figure 17: prior and posterior predictive checks

Ruta Slivkaite
201805872@post.au.dk

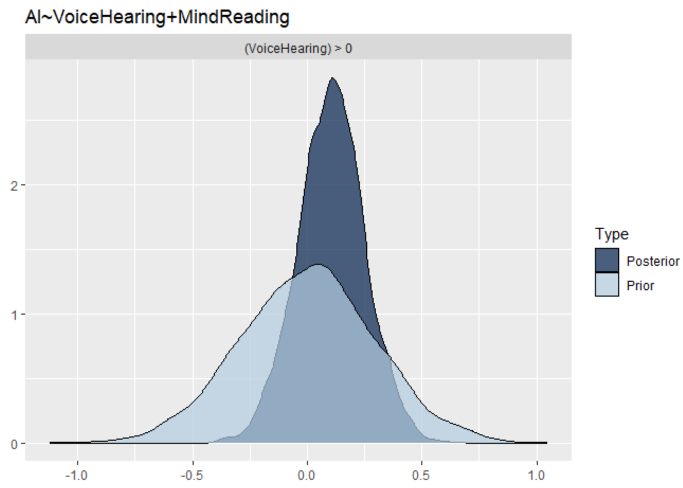


Figure 18

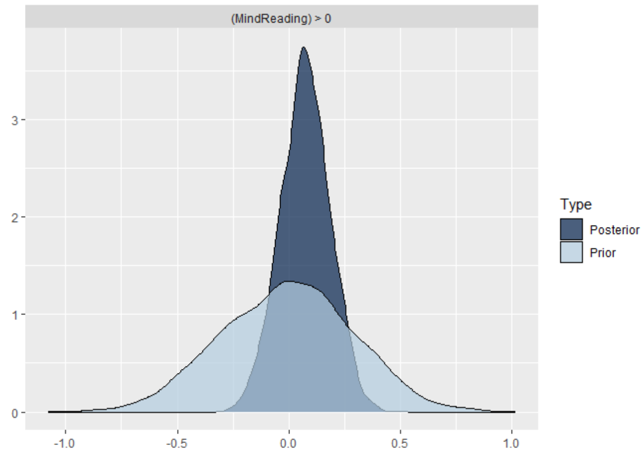


Figure 19

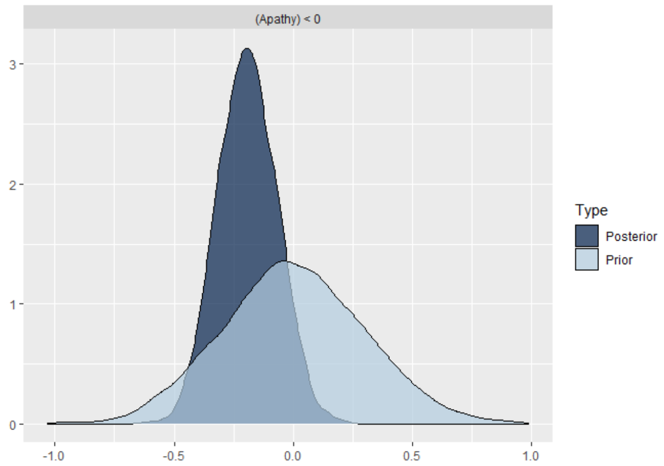


Figure 20

Figures 18-20: update check plots + hypothesis testing

Ruta Slivkaite
201805872@post.au.dk

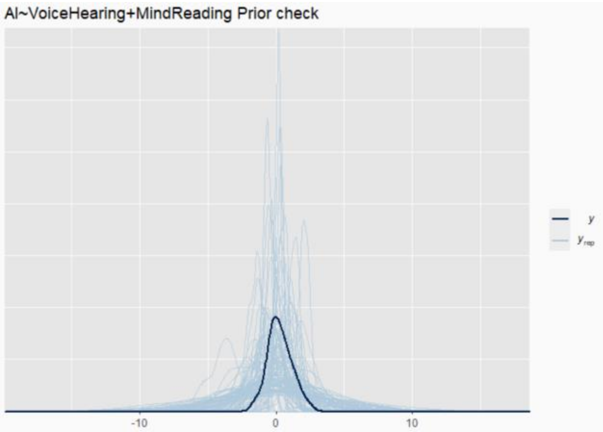


Figure 21 : prior predictive check check

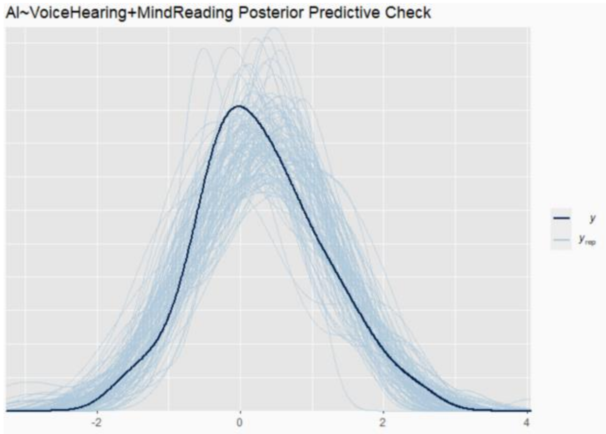


Figure 22: posterior predictive

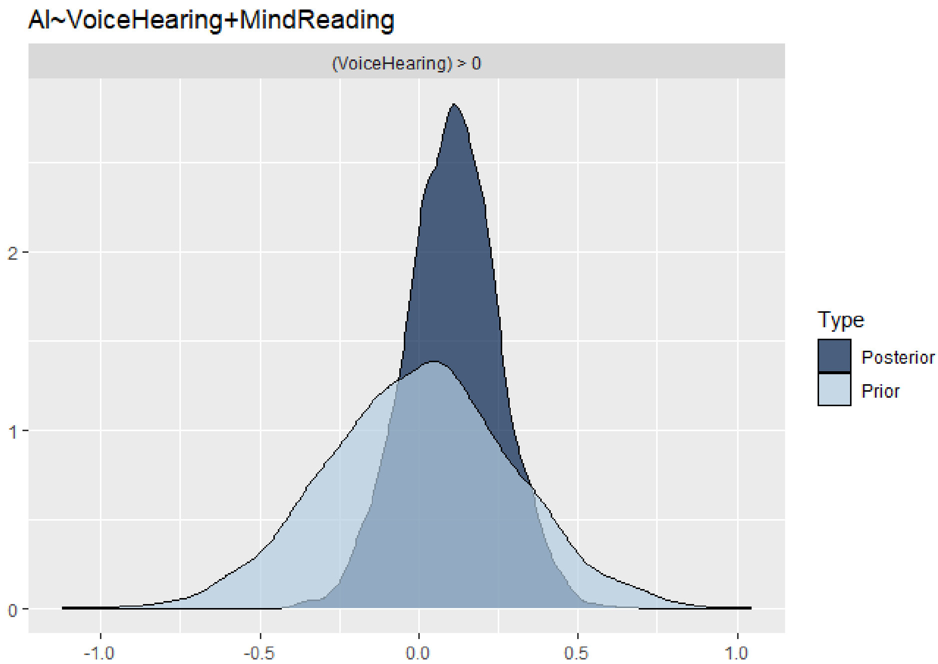


Figure 23 : update check plot + hypothesis testing

Ruta Slivkaite
201805872@post.au.dk

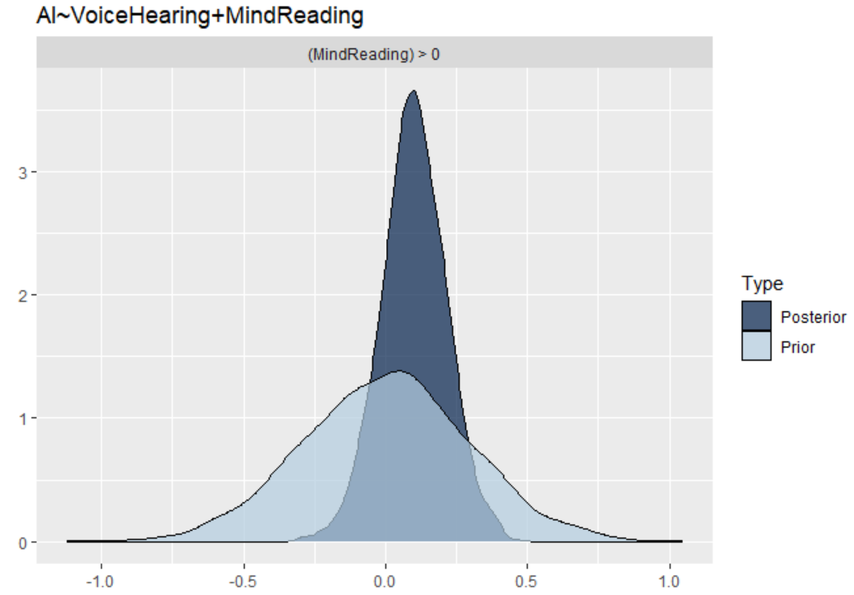
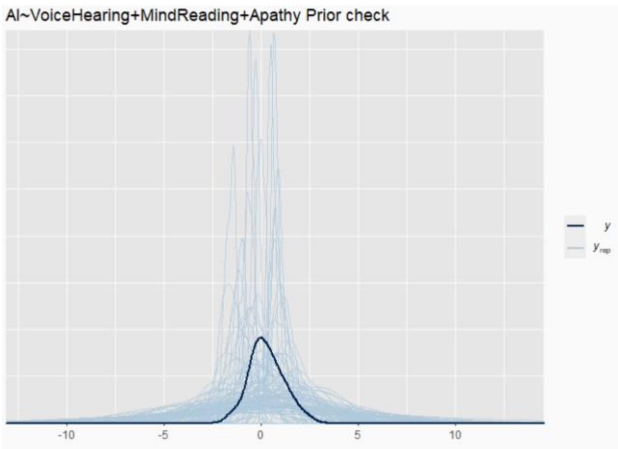


Figure 24 : update check plot + hypothesis testing



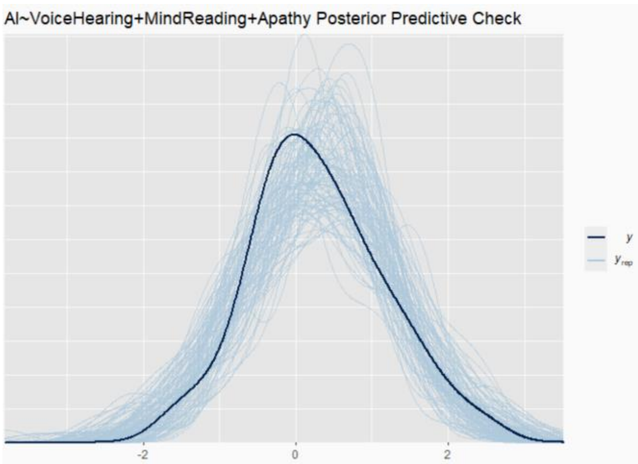


Figure 25 : prior predictive check check

Figure 26 : Posterior predictive

Ruta Slivkaite
201805872@post.au.dk

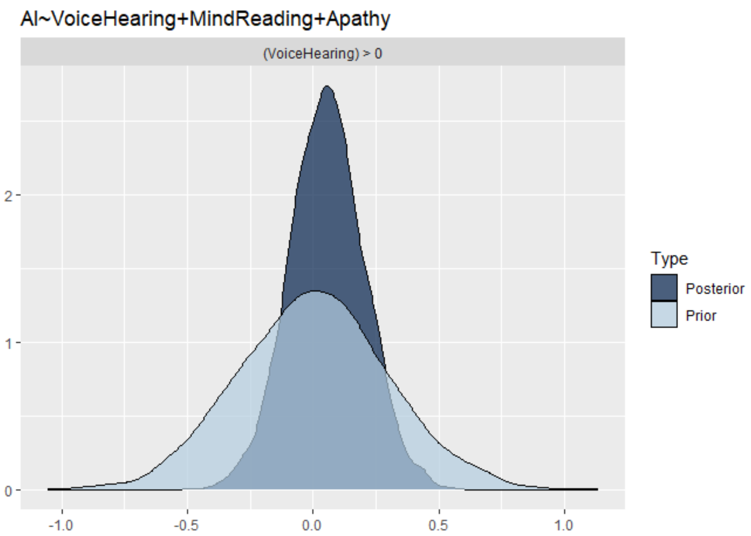


Figure 28

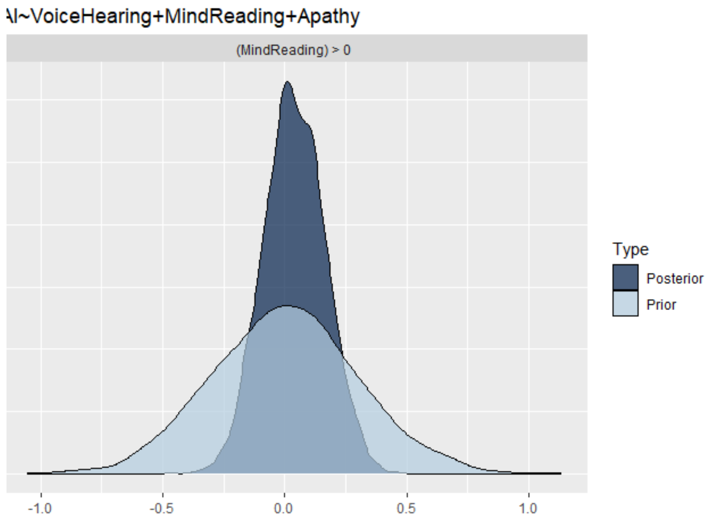


Figure 27

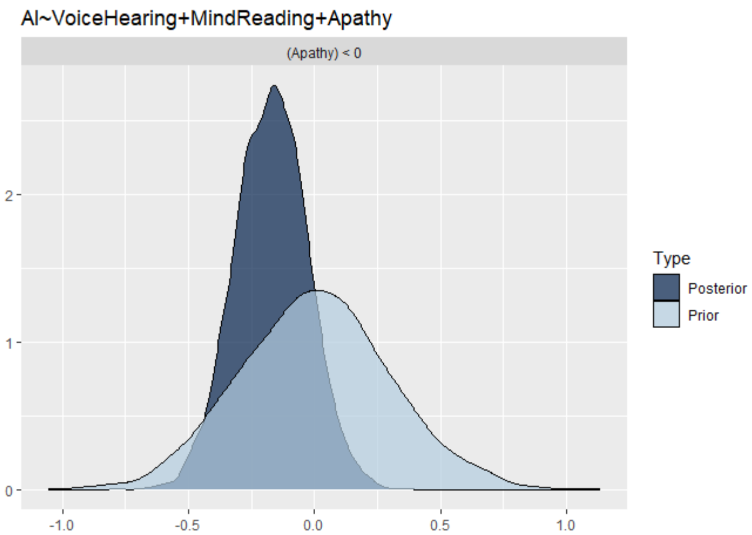


Figure 29

Figures 27-29 updating check plots + hypothesis testing

Ruta Slivkaite
201805872@post.au.dk

Assignment 4

Applying meta-analytic priors

Github link: https://github.com/PeterThramkrongart/CompMod4/blob/master/Assignment4.md

○

Introduction

In Bayesian modeling, choosing the right priors is, arguably, of the highest importance. There are multiple approaches to designing priors. Besides the most common one, a conservative prior choice, one can also employ meta-analytic priors to "stand on the shoulders of giants". Although it aptly reflects the cumulative nature of science, it is of vast importance that one remains critical. To shine a light on this matter, we compare a conservative set of priors to meta-analytic ones in the analysis of pitch variability in autistic and control subjects.

Both in the fields of social and psychological studies, one recurring concern is the generalizability of a sample to the entire targeted population. A possible remedy to this problem is the use of meta-analysis, as it compiles multiple heterogeneous samples into a single, more robust estimate. Therefore, it is more likely applicable to the entire population, including the sample of one's study. On the other hand, the quality of the meta-analysis is derived from the quality of the experiments included. Following the same rationale, compromised studies necessarily result in a compromised meta-analysis. However, there are other pitfalls along the way. When using meta-analytic priors, one must be aware of the possible cultural, national, or linguistic differences, as the effects might not be cross-culturally consistent. In theory, if the meta-analysis is rigorous, this approach is an adequate starting point for further studies, utilizing meta-analytic priors.

To see whether this theory applies to practice, we decided to put it to a test. We estimated pitch variability of autists in two studies – one American and one Danish. It is hypothesized that autistic people have a largely increased pitch and pitch variability, but we also want to test whether this difference was dependent on language. We employ Bayesian linear regression with varying effects and two priors. Either a conservative choice of our own or one based on the meta-analysis of 41 studies.

Hypotheses

The hypothesis we aim to evaluate in this paper is:

H1- Children with Autism Spectrum Disorder will have different pitch variability than their Typically Developing peers, however the effect may vary across languages.

Ruta Slivkaite
201805872@post.au.dk

Modeling with conservative priors

As we have data from two studies of different languages, we need to appropriately account for it. Therefore, we predict pitch variability from an interaction of diagnosis with language and varying effects for every participant. To know the relevance of the interaction, we further compared the interaction model with a baseline model. The formulas go as follows:

```
Baseline model: Pitch variability ~ 1 + Diagnosis + (1 | Participant ID))

Interaction model: Pitch variability ~ 0 + Language + Diagnosis:Language +  (1 |
Participant ID))
```

Designing conservative priors

In the process of designing the priors, we remained rather skeptical to prevent overfitting. As the data is standardized to Cohen's d, all the variables are automatically zero-centered. The distribution of pitch variability is roughly gaussian with some skew towards lower values. A log-transformation skews the data to a similar degree towards higher values. Therefore, we decided to use Gaussian priors as transforming variables in Cohen's d form to log distribution is above our abilities.

Choosing identical priors of intercepts for both languages reflects our skepticism in the hypothesis, since we do not presuppose a language difference. Nonetheless, by choosing not too restricting priors, we allow the difference to manifest, if there is any. The prior for interaction estimates are zero-centered with a small standard deviation as, in general, the effects in psychology are modest. Furthermore, by setting the same prior for interaction estimates of both languages, we let the data 'talk'. Further prior concerns are the varying effects of participants, which is again zero-centered with a low standard deviation, as we do not expect our participants to vary dramatically. Lastly, a prior for the model's expected average error, sigma, is set to 0.5 as the standard deviation of the outcome variable, pitch variability, is 1 and we do not expect the model to do worse.

$$\beta_{Languagedk} \sim normal\ (0,.1)$$
$$\beta_{Languageus} \sim normal\ (0,.1)$$
$$\beta_{Languagedk:DiagnosisTD} \sim normal(0,.1)$$
$$\beta_{Languageus:DiagnosisTD} \sim normal(0,.1)$$
$$\alpha \sim normal(0,.1), class\ =\ sd$$
$$\sigma \sim normal(.5,.0), class\ =\ sigma$$

Modeling with meta-analytic priors

Another approach for the analysis of our study's data is to employ meta-analytic priors. This is a sensible alternative to conservative priors since these priors are informed by past studies, which enables us to be more confident when, for instance, we are setting the standard deviations. The

Ruta Slivkaite
201805872@post.au.dk

model thereby has more relevant information to start with and any effect detected will generally be more robust.

Meta-analysis

To generate the exact values for the meta-analytic priors, we analyzed 41 studies concerning pitch variability in ASD and controls. This analysis was carried out using the formula below, which estimates the effect size of the autistic diagnosis on pitch variability along with its error term which concerns the measurement uncertainty. Also, we account for population by modeling random intercepts as the same population has been used multiple times.

```
EffectSize | se(StandardError) ~ 1 + (1 | Population)
```

By running this formula using multivariate Bayesian linear modeling and normally distributed priors for the intercept (0,1) and standard deviation (0,3), we get the output, described in Table 1. The mean of the effect size together with the population are used to create informed meta-analytic priors in the following section.

Variable	Estimate	Est. Error	Lower 95% CI	Upper 95% CI
Effect size	-0.43	0.09	-0.60	-0.25
Population variance	0.31	0.09	0.12	0.50

Table 1 - Output of the meta-analysis model

Designing meta-analytic priors

When rerunning the interaction model previously defined in the section regarding conservative priors, we now utilize the information from the meta-analysis to direct our model.

In selecting priors for the interaction model, the first two priors indicate the baselines for the languages: one for Danish, one for English. The priors for interaction effects were based on the effect size of the meta-analysis, while the average expected prediction error of the model was given by variation of populations in the meta-analysis. We, again, decide to utilize Gaussian distributed priors. The formula for the priors is given:

$$\beta_{Languagedk} \sim normal\ (.2, .3)$$
$$\beta_{Languageus} \sim normal\ (.2, .3)$$
$$\beta_{Languagedk:DiagnosisTD} \sim normal(-0.43, .1)$$
$$\beta_{Languageus:DiagnosisTD} \sim normal(-0.43, .1)$$
$$\alpha \sim normal(0, .1), class\ =\ sd$$
$$\sigma \sim normal(.32, .1), class\ =\ sigma$$

Ruta Slivkaite
201805872@post.au.dk

[1]

We repeated the procedure from the conservative priors using the meta-analytic priors and built the same two models as defined previously with pitch variability as the dependent variable and participant ID as a random effect. Furthermore, the baseline model uses diagnosis as a predictor, while the complex model models each language with separate intercepts and slopes dependent on diagnosis.

```
Baseline model: Pitch variability ~ 1 + Diagnosis + (1 | Participant ID))

Interaction model: Pitch variability ~ 0 + Language + Diagnosis:Language +  (1 |
Participant ID))
```

The results from running the multivariate regression model on the study data using the meta-analytic priors stated above are described in the section below. These results are compared to those obtained from the analysis using conservative priors. We wish to elucidate the discrepancies between the two approaches to setting priors and discuss the applicability as well as the pitfalls of both methods.

Comparative discussion of methods and results

Predictive checks

Both model priors seem reasonable and within range when prior predictive checks are performed. As expected, they do not quite catch the skew of the data and have a center peak around zero. (see Figure 3 and 4 in Appendix). Similarly, both models describe the data quite well when posterior predictive checks are performed (see Figure 3 and 4 in Appendix). Again, the models don't perfectly catch the skew of the data.

Results and estimates

According to the model employing the conservative priors, the mean pitch variability in ASD is much higher in American-English (b=0.43, CI´s=0.31, 0.55) compared to Danish (b= -0.14, CI´s= -0.26, -0.01) individuals. Looking at the estimates of the model with informed priors, the mean pitch variability in ASDs is, again, much higher in American English (b= 0.77, CI´s= 0.62, 0.93) compared to Danish language (b=-0.05, CI´s= -0.20, 0.10). Additionally, the employment of informed priors has led to an

[1] We have unfortunately introduced a small human error in our code by defining the sigma prior value as 0.32 instead of 0.31. Considering the rather small size of the error of our typo and the extended running time of the code as well as the amount of spillover effects it would have on our other values, we have decided to continue using sigma = 0.32, .1.

Ruta Slivkaite
201805872@post.au.dk

increase in the mean pitch value of ASDs in both languages compared to the skeptical prior estimates. The visual representation of the diagnosis effect on pitch variability across both languages is depicted in Figure 1 and Figure 2 and is discussed in more detail in the hypothesis testing section.

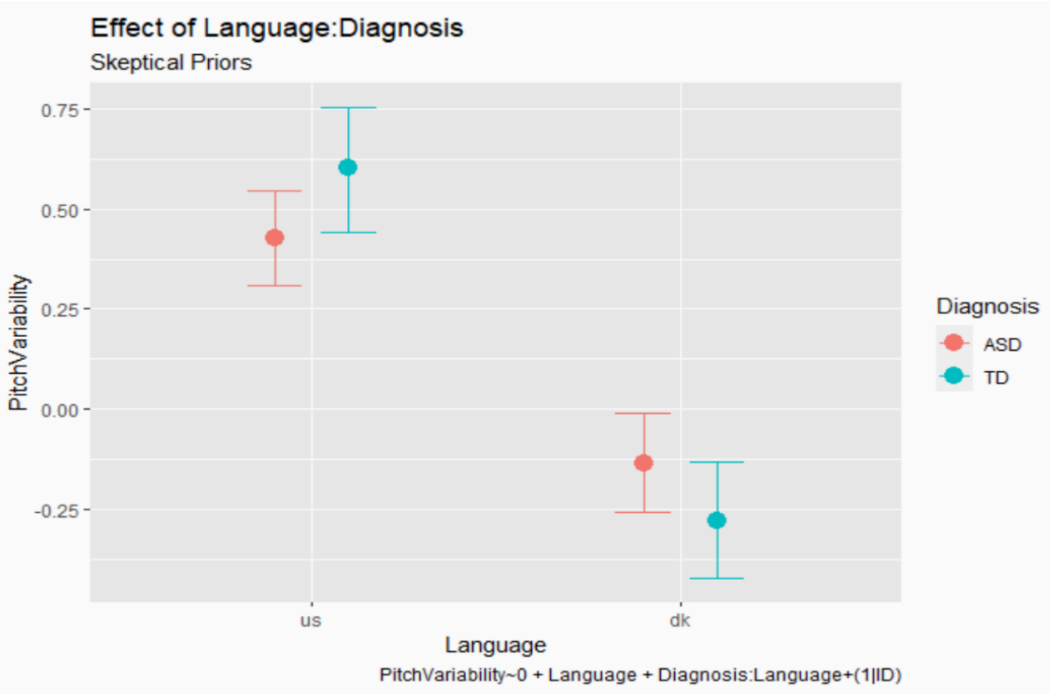


Figure 1 - Pitch variability in two languages by diagnosis using skeptical prior

Ruta Slivkaite
201805872@post.au.dk

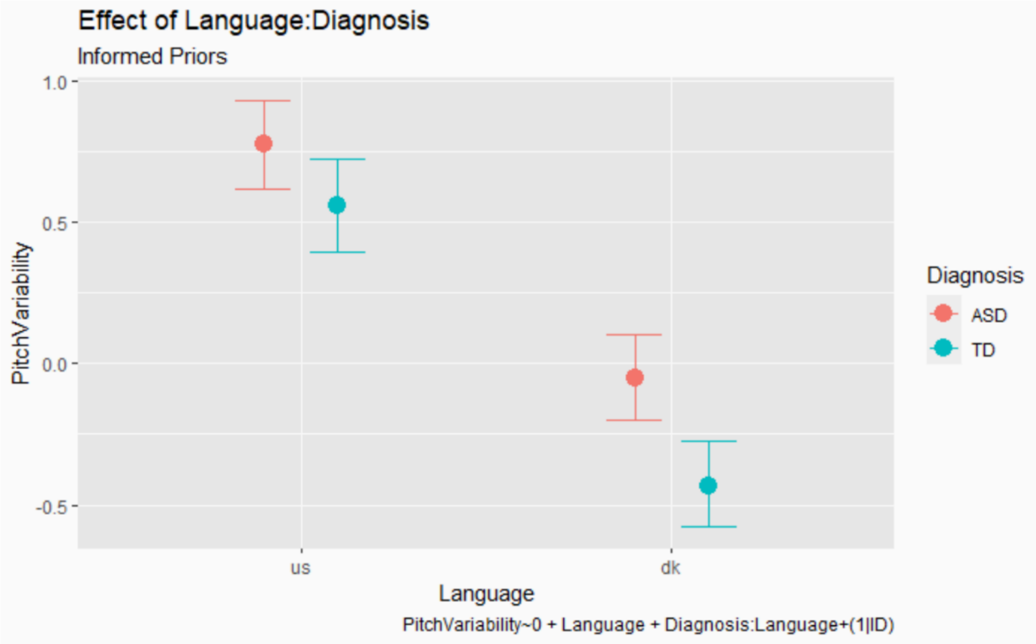


Figure 2 - Pitch variability in two languages by diagnosis using informed prior

	Estimate	Est. Error	Lower 95% CI	Upper 95% CI
Model with skeptical prior				
Languagedk	-0.14	0.06	-0.26	-0.01
Languageus	0.43	0.06	0.31	0.55
Languagedk:DiagnosisTD	-0.14	0.08	-0.29	0.01
Languageus:DiagnosisTD	0.17	0.08	0.03	0.33
Model with informed prior				
Languagedk	-0.05	0.08	-0.20	0.10
Languageus	0.77	0.08	0.62	0.93
Languagedk:DiagnosisTD	-0.38	0.08	-0.54	-0.23
Languageus:DiagnosisTD	-0.22	0.08	-0.37	-0.06

Table 2 - Results of the models with both skeptical and informed priors

Hypotheses testing

Our prediction that the mean pitch variability in individuals with Autism Spectrum Disorder will differ by language as well as in comparison to TD children has been confirmed. When modeling with skeptical priors, Danish neurotypical children display a slightly lower mean pitch variability than their ASD peers (b=-0.14, CI´s=-0.27, -0.01, ER=29.77). In contrast, American neurotypical children show a slightly higher mean pitch variability (b=0.17, CI´s=0.05, 0.31, ER=0.01). However, this effect disappears when we employ our informed prior, with both languages showing a lower mean pitch

Ruta Slivkaite
201805872@post.au.dk

variability in TD patients. Danish neurotypical children exhibit a much more credible lower mean pitch variability (b=-0.38, CI´s=-0.51, -0.25, ER=Inf) from ASDs than their American peers (b=-0.22, CI´s=-0.35, -0.08, ER=284.71). However, it needs to be pointed out that informed priors for the language-diagnosis interaction was set with mean -0.43 and very small standard deviation of 0.1. That inevitably leads to exclusively negative estimates. Therefore, testing whether the posterior samples of the interactions are smaller than zero is not informative as the informative prior does not allow this. Hence, instead of consulting evidence ratio (written in italics), which were included for completeness, in this case, one' attention should be focused at credible intervals instead.

When we test the hypothesis whether the pitch variability in ASDs in Danish language is lower than their peers in English American language we get it confirmed with both priors, where the difference is bigger and more credible using skeptical prior (b=-0.32, CI´s= -0.5, -0.13, ER=399) compared to informed prior (b=-0-17, CI´s=-0.36, 0.02, ER=12.42).

The models had no divergences, a Rhat of 1 and Effective sample sizes well above 200 for both Bulk and Tail.

	Estimate	**Est. Error**	**Lower 95% CI**	**Upper 95% CI**	**Evid. Ratio**
Model with skeptical prior					
Languagedk:DiagnosisTD < 0	-0.14	0.08	-0.27	-0.01	29.77
Languageus:DiagnosisTD < 0	0.17	0.08	0.05	0.31	0.01
Languagedk:DiagnosisTD < Languageus:DiagnosisTD	-0.32	0.11	-0.5	-0.13	399
Model with informed prior					
Languagedk:DiagnosisTD < 0	-0.38	0.08	-0.51	-0.25	*Inf*
Languageus:DiagnosisTD < 0	-0.22	0.08	-0.35	-0.08	*284.71*
Languagedk:DiagnosisTD < Languageus:DiagnosisTD	-0.17	0.11	-0.36	0.02	12.42

Table 3 - Hypotheses testing of the models with both skeptical and informed priors

Comparing methodological approaches

Model	**Weight**
With skeptical prior	0.413
With informed prior	0.587

Table 4 - Leave-one-out cross-validation of the models with both skeptical and informed priors

In a small world environment where our two models with different priors, skeptical and informed, respectively, are the only possible models, the leave-one-out cross-validation predicts that there is a 59% probability that the model with informed priors is the best model (see *Table 4*). However, we

Ruta Slivkaite
201805872@post.au.dk

should not blindly assume that the informed priors offer the best, most robust results every time, because of the possible drawbacks associated with meta-analytic priors, which were mentioned previously in the introduction. Furthermore, in this specific example, the LOOCV weight is quite close to chance, which also indicates that we don't have sufficient evidence to be able to confidently choose one model over the other. On the other hand, as long as we are aware of the reliability of our sources, which are used for setting the informed prior, it is considered an applicable practice to aptly make use of the accumulation of scientific knowledge.

Using skeptical priors, the effect of diagnosis differs across languages. Observing the American-English subjects, we can detect lower pitch variability in ASD patients, whereas when looking at the Danish subjects, ASD patients appear to have higher pitch variability than neurotypical controls. However, when the meta-analytic (informed) priors are employed, the effect of diagnosis is consistent across languages. ASD subjects were found to have higher pitch variability regardless of their spoken language.

There are several possible interpretations of what we witness in our case. One explanation could be that the idiosyncrasies of the samples from the two studies, which were used to motivate our skeptical priors, might be larger than usual. When informed priors are used, this increased sample variance was corrected, and the "true" effect thus emerged. Another explanation could be that the effect of ASD diagnosis on pitch variability is, in fact, not consistent across languages. As the informed priors are based predominantly on the English-speaking population (26 studies out of 41 in the meta-analysis), the "true" effect in our analysis could be therefore falsely corrected, when relying on meta-analytic priors.

Given the specific context of our study, it is difficult to generalize the efficiency of using meta-analytic priors. However, there is still a lesson to be learned. Accounting for cultural and language differences is certainly as important of an aspect when employing meta-analytic priors as it is when we are using skeptical ones. Conceptually, the usage of meta-analytic priors is well-reasoned, and attempts of its implementation should continue, however, we need to be sure that we are standing on the shoulders of the right giant.

In terms of our research question, we cannot draw any resolute conclusions, since it's not obvious which model is best. If we want to investigate the matter further, we suggest redoing the meta-analysis, defining that the effect of diagnosis on pitch variability should be analyzed individually for each language in the meta-analysis. In the case of our study, when we defined the model with informed priors, we set the same mean and standard deviation for the effect of language on diagnosis. If we truly want to utilize the cumulative nature of science, it seems natural to take language into account from the start. Otherwise we might be restricting our outcome to an unnecessary degree.

Ruta Slivkaite
201805872@post.au.dk

Conclusion

To conclude, when evaluating which priors to use, conservative and meta-analytic priors can generate different results. Our hypothesis H1 is validated when we model our data with both skeptical and meta-analytic priors. The directions of the effects, however, change depending on the chosen priors. When using skeptical priors, the model predicts a decrease in pitch variability within American ASD children and increase pitch variability within Danish ASD children. Although, this effect changes when employing informed priors, as this model predicts an increase in pitch variability regardless of the language.

Evidently, it is important to substantiate the theoretical reasoning behind choice of prior, and when in doubt, it could be relevant to assess what effect the use of conservative versus informed priors contributes to the analysis. In theory, meta-analytic priors are preferred, since they prioritize robustness as well as capitalize on years of hard work from other researchers. In practice, however, they are susceptible to human error. Furthermore, they might impress a false sense of confidence if applied inconsiderately. On the other hand, the informed priors prevent us from being too rash when encountering new evidence, which conflicts from the accepted paradigms. Furthermore, because of the immense amount of past research we can assume non-zero prior values for the mean of the effect size, enabling the model to gain more confidence than achieved by applying skeptical priors, given the new evidence is consistent with the meta-analytic effects. The reliance on other researcher's data using meta-analytic priors is thus a double-edged sword, which generates reliable results if wielded thoughtfully.

Appendix

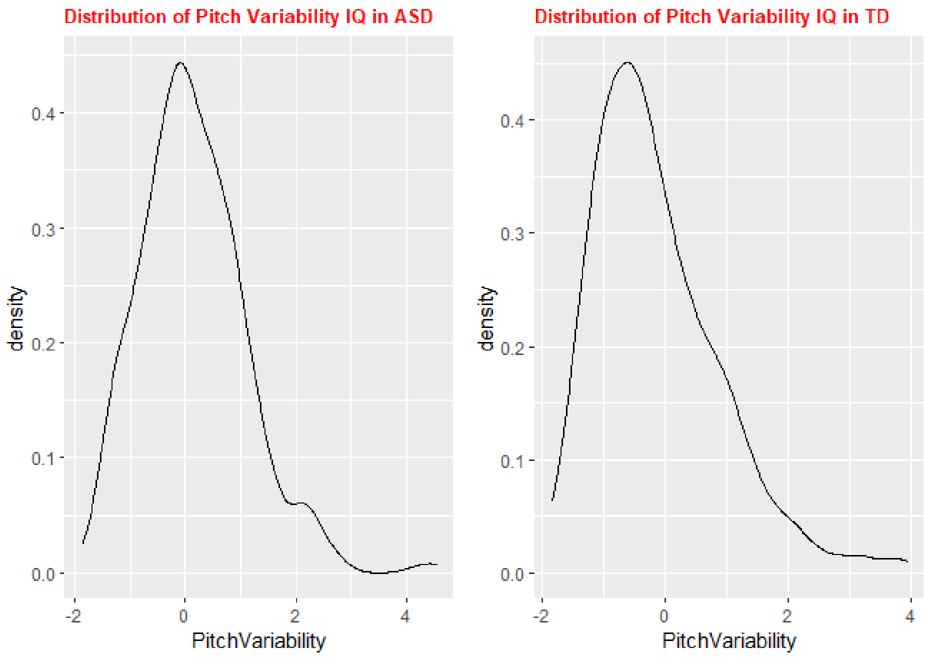


Figure 3 - Distribution of pitch variability in ASD and TD

Ruta Slivkaite
201805872@post.au.dk

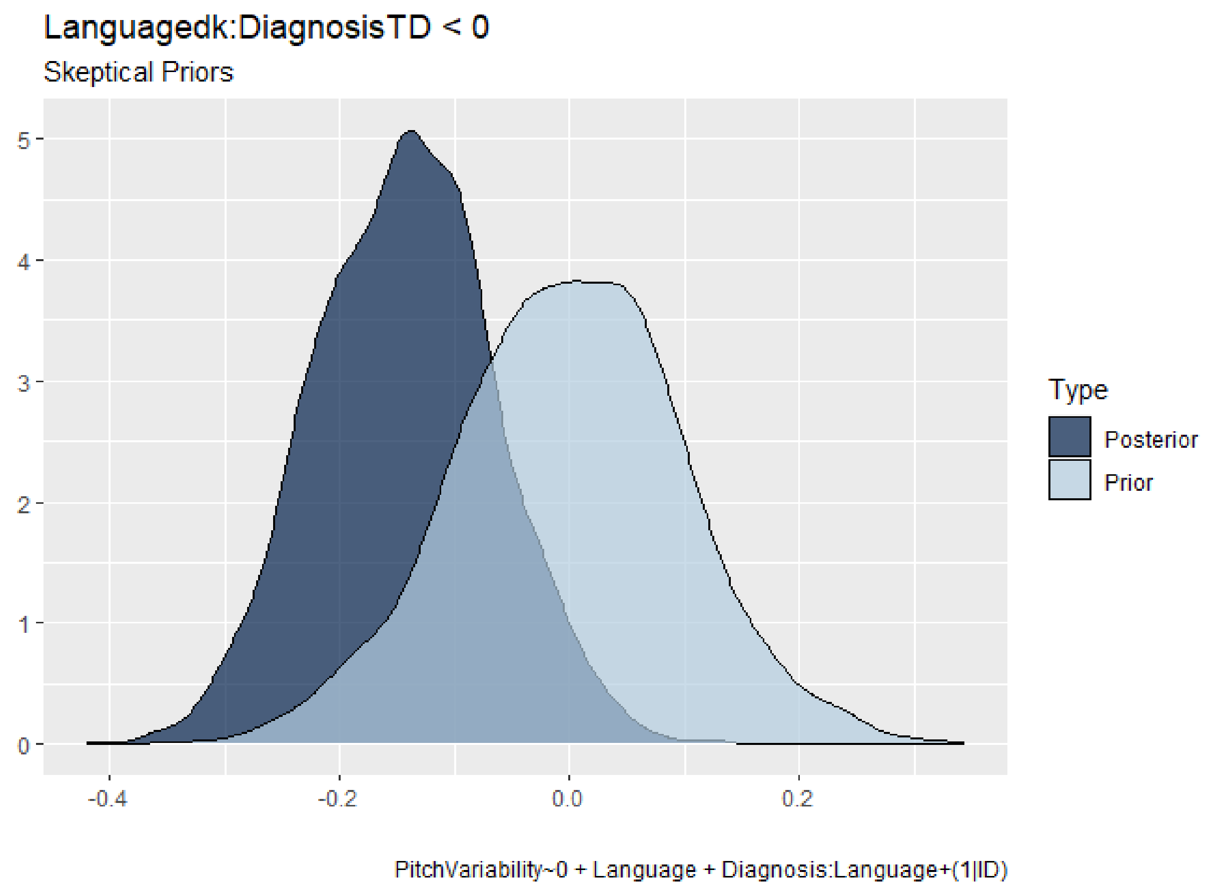


Figure 4 - Hypothesis testing with skeptical priors

Ruta Slivkaite
201805872@post.au.dk

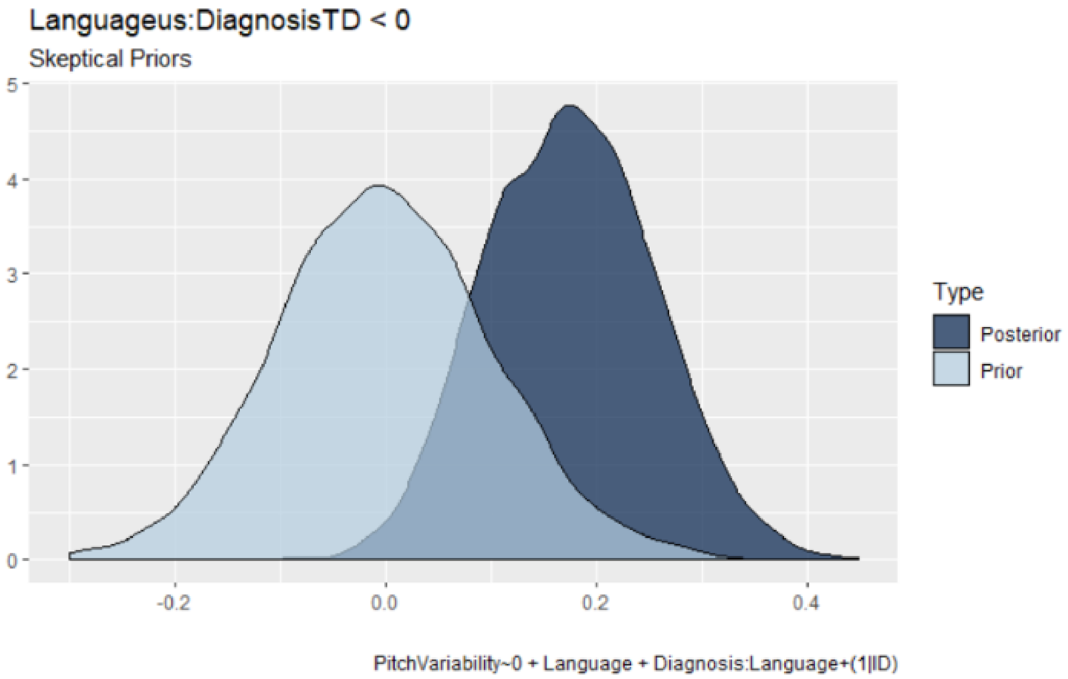


Figure 5- Hypothesis testing with skeptical priors

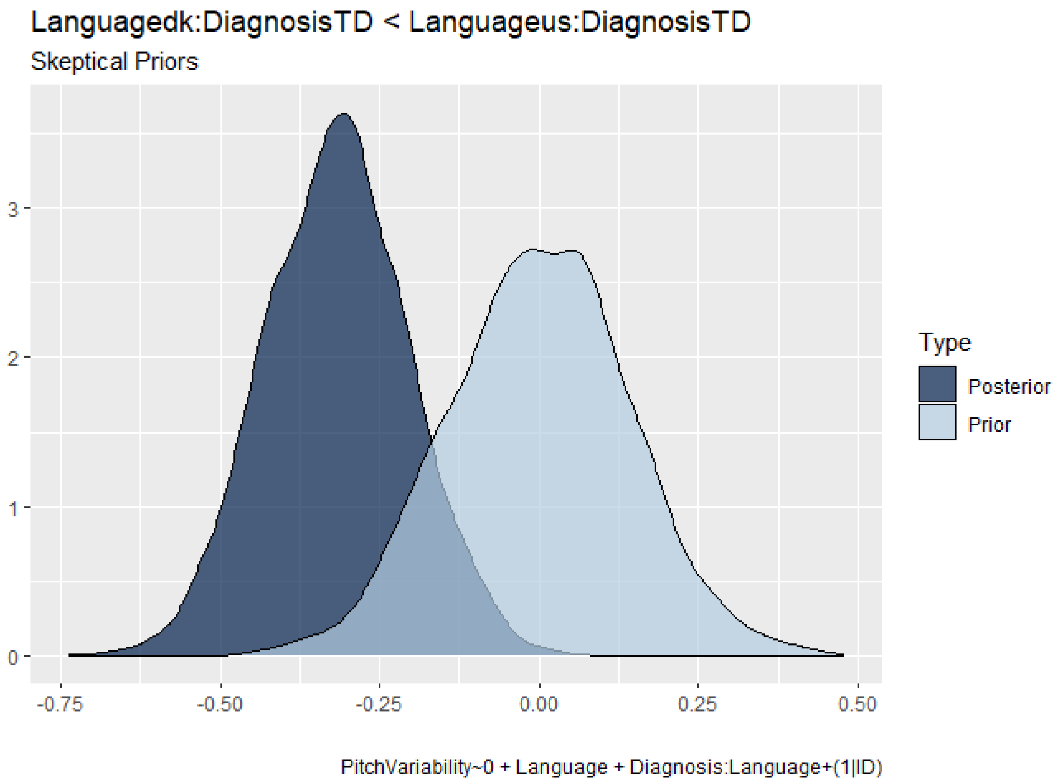


Figure 6 - Hypothesis testing with skeptical priors

 Ruta Slivkaite
201805872@post.au.dk

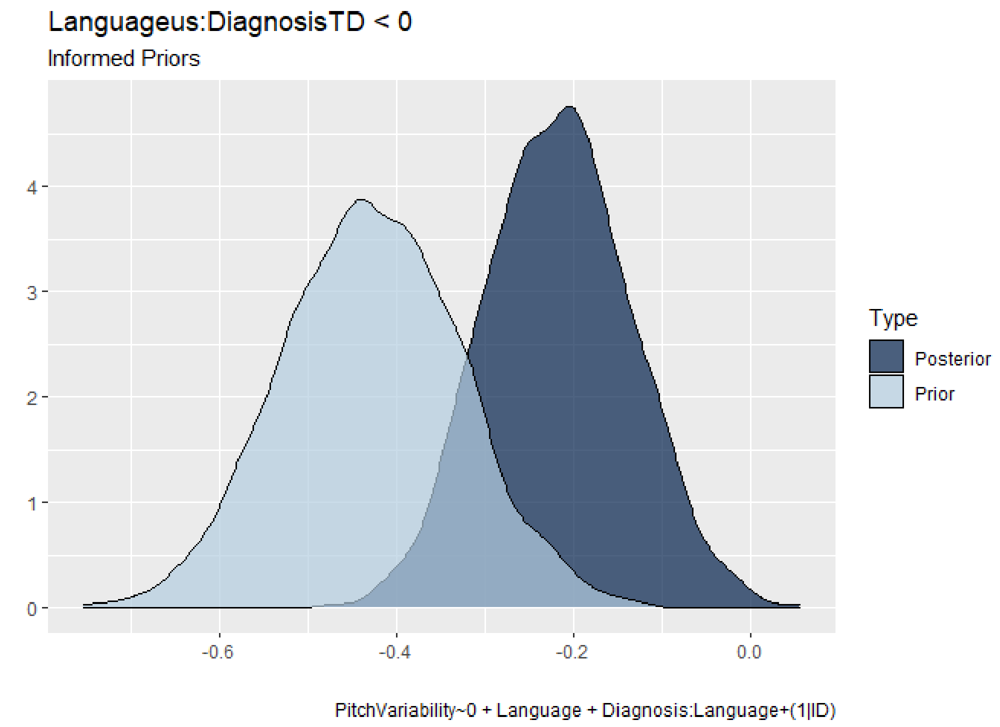


Figure 7 - Hypothesis testing with informed priors

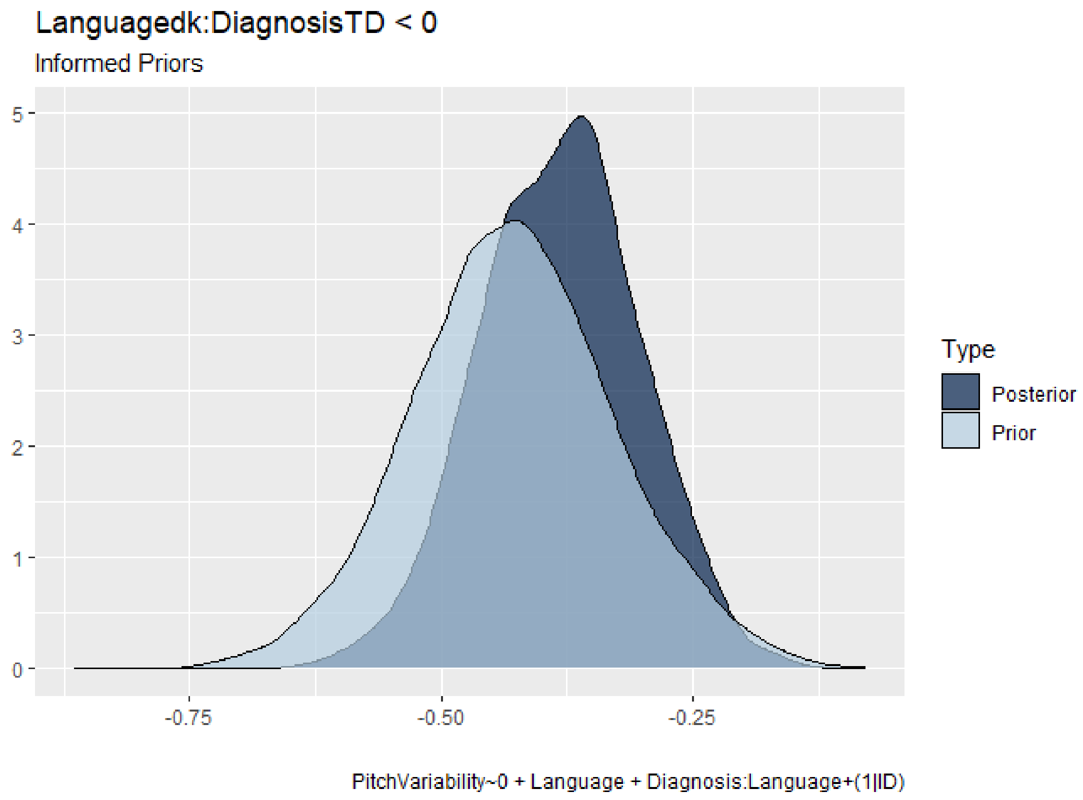


Figure 8 - Hypothesis testing with informed priors

Ruta Slivkaite
201805872@post.au.dk

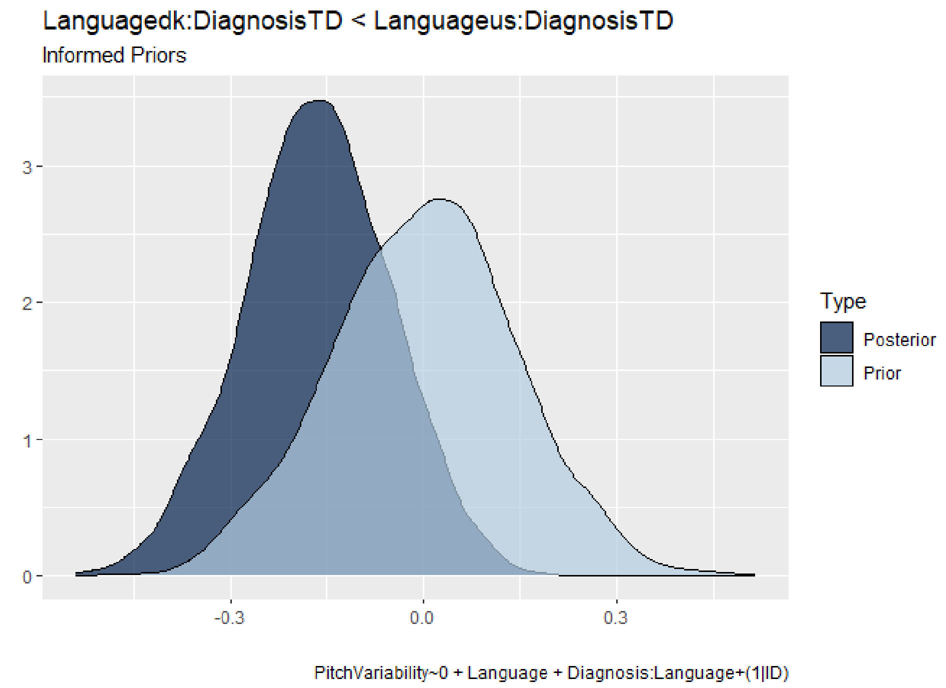


Figure 9 - Hypothesis testing with informed priors

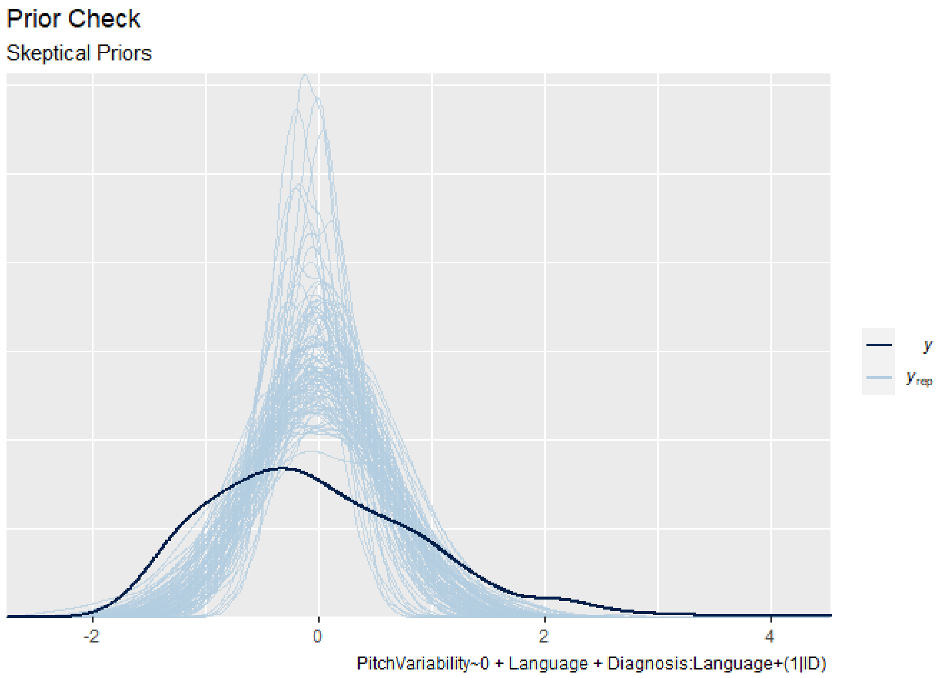


Figure 10 - Prior Predictive Check with skeptical priors

Ruta Slivkaite
201805872@post.au.dk

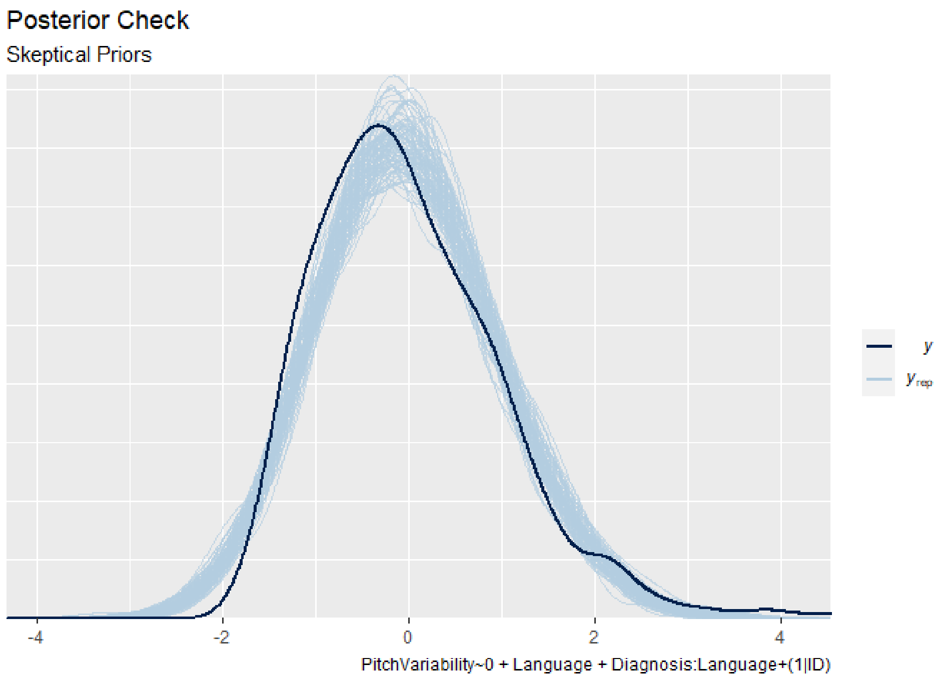


Figure 11 - Posterior Predictive Check with skeptical priors

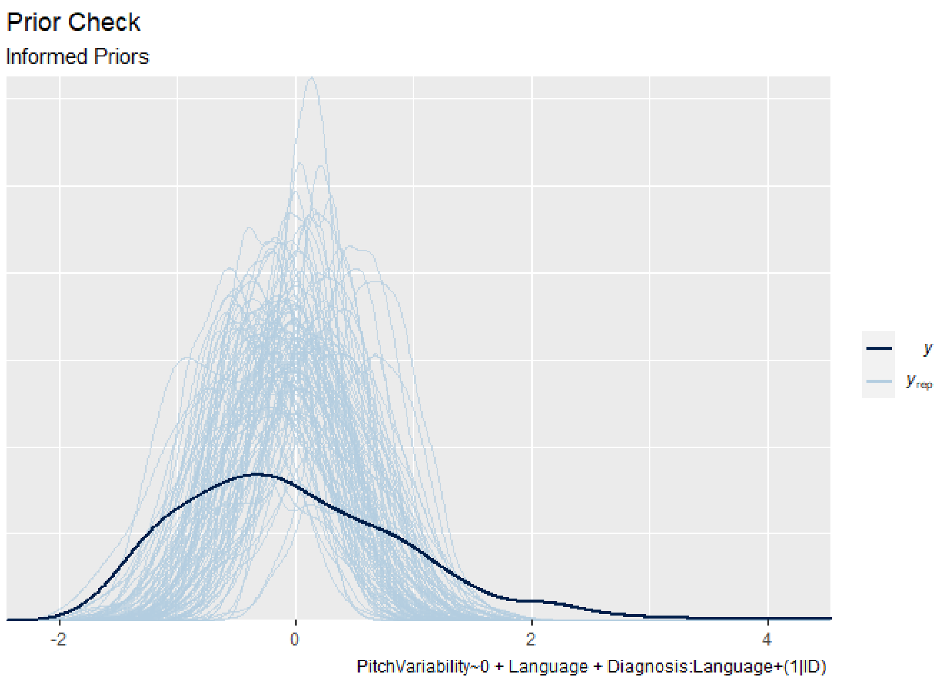


Figure 12 - Prior Predictive Check with informed priors

Ruta Slivkaite
201805872@post.au.dk

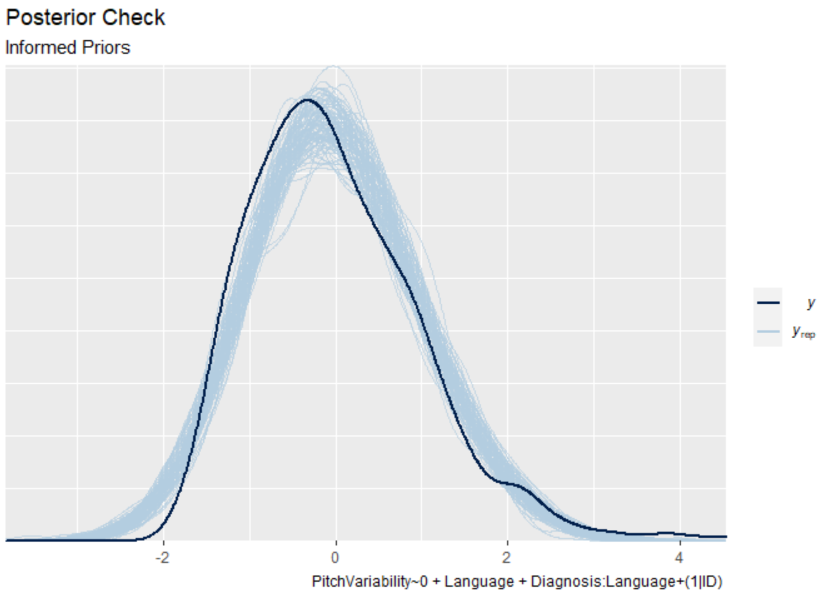


Figure 13 - Posterior Predictive Check with informed priors

Ruta Slivkaite
201805872@post.au.dk

Assignment 5

Agent based model of the role of social support in the development of Major Depressive Disorder

Github: https://github.com/Rutatu/CompMod4-ABM/blob/master/My-ABM.md

Introduction

Questions and hypotheses

Major Depressive Disorder (MDD) is a complex mood disorder with many causing factors at play. However, scientists formed a general theory to explain the causes of psychiatric disorders and their course over time, including MDD, called the stress--vulnerability model. According to the model, three critical factors are responsible for the development of a psychiatric disorder and its course over time: biological vulnerability, stress and protective factors. In this paper I have decided to employ this more generic model to build an agent based model (ABM) to explore these questions: 1) how stressful life events by increasing the distress level could lead to an increased likelihood of MDD onset in high and low vulnerability groups of people; 2) how and to what extent two distinct kinds of social support: perceived and received social supports, interact and act as a buffer against distress thus altering the likelihood of MDD onset in high and low vulnerability groups of people.

Hypotheses:

H1: Stressful life events would lead to a considerably increased likelihood of MDD onset in a high vulnerability group compared to a low vulnerability group despite received support scenario.

Ruta Slivkaite
201805872@post.au.dk

H2: Higher perceived support level will buffer distress and will lower the likelihood of MDD onset the most compared to lower perceived support levels.

H3: Appropriate received support scenario will buffer distress and will lower the likelihood of MDD onset the most compared to other scenarios.

Agent based model

The ABM was described and formalized following the ODD (Overview, Design concepts, Details) protocol (Grimm et al., 2010), focusing only on those elements that are relevant to this specific ABM.

Overview

Purpose

The model addresses the role of distress caused by stressful life events and its impact on the MDD onset while taking into account the biological vulnerability of the population. The model explores the relationship between the social support, which acts as a protective mechanism against distress, and the likelihood of developing an episode of MDD.

Entities, state variables and scales

The model entities include agents representing individuals who form two populations: low and high vulnerability populations. Each population consists of 150 individuals creating a total population of 300 individuals or agents. Each agent is characterized by these state variables or attributes: *identity number, perceived support level, vulnerability group, distress level, MDD onset, appropriate received support count, inappropriate received support count (*see Table 1*).* The model's major components and their interaction can be seen in Figure 1.

Ruta Slivkaite
201805872@post.au.dk

State variable	Scale
Identity number	1-300
Perceived support level	A positive number on a continuous scale with a minimum values of 0 and a maximum value of 10
Vulnerability group	Low vulnerability (L) or High vulnerability (H)
Distress level	A positive number on a continuous scale starting from 0 with no end boundary
MDD onset	Either 1 or 0 (gets updated after the simulation)
Appropriate Received support count	Positive natural number from 0 to infinity (being updated throughout the simulation)
Inappropriate Received support count	Positive natural number from 0 to infinity (being updated throughout the simulation)

Table 1 - Agents´ state variables and their scales

Perceived support refers to one's subjective evaluation of the quality and trust in their connections which might be influenced by early developmental experiences and personality traits.

Distress level is an emotional constraint caused to an individual by specific events. The higher the value the more distress the individual is experiencing.

According to the literature, people can either be in a low or high *vulnerability group* based on their sensitivity to possible triggers or/and genetics. For the sake of simplicity, the creation of more individualistic vulnerability groups was avoided.

Ruta Slivkaite
201805872@post.au.dk

MDD is characterized by symptoms, which can be triggered by stressful life events and other factors at play. When a sufficient number of symptoms is present, the person is likely to experience an episode of MDD. In this model the symptoms were not explicitly modeled, instead, the implicit assumption, that a high enough level of distress would activate enough symptoms for the episode to happen, was made. Thus, *MDD onset* indicates whether an individual is likely to experience the first episode ever (1).

Received support refers to the reported utilization or exchange of support resources and is more likely to represent a situational factor that is sought or provided in response to stress whether it is asked or not asked support.

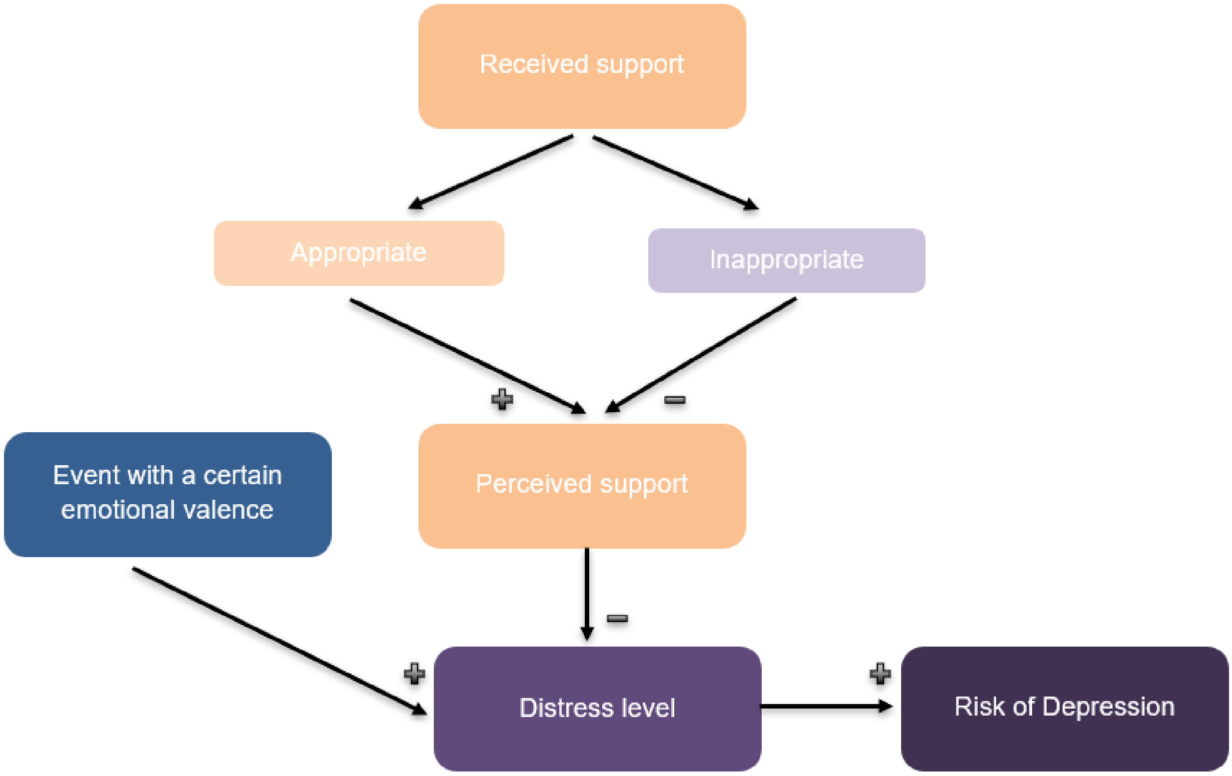


Figure 1 – ABM´s dynamics ("+" indicates that one component increases the other and "–" indicates that it decreases)

Ruta Slivkaite
201805872@post.au.dk

Environment

Environment or forces that drive the behavior and dynamics of every agent are represented by *stressful life events* that happen to the agents. The function creating the dataset of events with different emotional valence was created. The events are divided in four groups: *extremely negative*, *negative*, *moderate* and *low negativity* events characterized by *valence,* which is an emotional value of an event, and *impact,* which is a set value indicating how much distress event elicits based on its valence (see Table 2) . The valence is represented on a continuous scale from -13 to 0 and for each event was generated using uniform random numbers, for the specific intervals see the Table 2 below. The intervals, valence scale and impact value were chosen based on both the literature and subjective assumptions that could reflect real life to some extent.

Event	Valence	Impact
extremely negative	-13 to -12	14
negative	-12 to -8	7
moderate	-8 to -4	3
low negativity	-4 to 0	1

Table 2 - Stressful life events with their valence intervals and impact values

The simulation was run using a dataset of 15 events in total: *extremely negative (1 event)*, *negative (2 events)*, *moderate (4 events)* and *low negativity (8 events)*

The idea behind choosing this set of events was based on the literature indicating that one extremely negative event in addition to some other less negative events can cause MDD onset. This

Ruta Slivkaite
201805872@post.au.dk

scenario excluded an opportunity of positive events happening besides the appropriate received support in order to see the protective influence of social support more clearly.

The length of a time step is unspecified but is represented by the occurrence of one event at a time, which could represent a day, a week or a month in real life.

Process overview and scheduling

The following actions are executed, in this order, once per time step/event:

- Events, generated by the function, happen one at a time and each of them increases the distress level of every agent to the same extent which corresponds to its *impact* value, which makes no one totally immune to the stress (see Table 2 above).
- One of the three received support scenarios happen randomly to each agent during each time step/event with a 33.33% chance: a) agent receives appropriate support, b) agent receives inappropriate support, c) agent does not receive any support (see submodel *received support*).
- The number of received appropriate or inappropriate support is being counted for each agent and saved to the dataset after each event.
- In turn, received support increases/decreases perceived support level depending on its appropriateness and the value of perceived support is updated (see submodel *perceived support update*).
- The individual buffer is calculated which indicates how much perceived support buffers against distress and the value is subtracted from the current distress level of an agent (see submodel *individual buffer*).
- An agent is either assigned an onset of MDD or not (see submodels *low vulnerability* and *high vulnerability*)

Ruta Slivkaite
201805872@post.au.dk

Design concepts

This section describes and justifies the key design elements that were used to design the ABM. In this case, only relevant elements were used and for the full list of possible elements see Figure 2 and Figure 3 *in the Appendix.* While mentioning these design principles is not needed to replicate a model, this section may be crucial to interpreting the output of a model.

Emergence

The key outcomes of the model are distress level and the presence/absence of the onset of MDD. These outcomes emerge mainly from the interaction between stressful life events and perceived support level. The outcome is not fully emergent in the way that the interactions producing the outcome were clearly defined by the ABM and very surprising results were not expected.

Adaptation

The adaptive behavior is reflected by the perceived support acting as a protective mechanism, which decreases distress level, and could correspond to the individual´s events´ appraisal mechanism in real life***.***

Objectives

There is no explicit objective that the agents exhibit in the ABM. Nevertheless, the idea behind the protection against distress that perceived support together with received support offers could be understood as an implicit objective.

Interaction

The main interactions in the ABM are between the events and distress level of the agents and both kinds of support. There is no interaction between agents itself.

Stochasticity

The stochasticity was used to determine whether agents receive appropriate, inappropriate or no social support. It was modelled so based on the assumption that received support is not always

Ruta Slivkaite
201805872@post.au.dk

present during each adverse life experience and depends on many other factors that are out of this ABM´s scope, thus it was decided to leave it for the random chance.

Collectives

The whole population of agents create two collectives: low and high vulnerability groups, which are defined prior to the start of simulation and do not change over time.

Observation

For the analysis of the ABM these measures were collected and used: perceived support level, distress level, MDD onset, appropriate received support count and inappropriate received support count. All the measures were being updated throughout the simulations and saved to the dataset.

Details

Initialization

The population of agents with their state variables was generated prior to the simulation and saved to a dataset. The same dataset was used for every new run of simulation. There were 300 agents in total, 150 in low vulnerability and 150 in high vulnerability groups. Initial mean perceived support value in the low vulnerability group was 5.0 (sd=2.9) while in the high vulnerability group it was 4.7 (sd=3.0). More details about all state variables can be found in the section *Entities, state variables and scales* and *Environment.*

There were four runs of simulation in total each corresponding to one of the received support scenarios. The changes of the scenarios were performed in the code of the simulation itself.

Ruta Slivkaite
201805872@post.au.dk

Submodels

Submodels describe and elaborate on the processes listed in the section *Process overview and scheduling*. There were five submodels listed:

- *Received support.* This submodel is performed after each event. Each agent gets assigned one of the three possible scenarios by drawing randomly from a uniform real distribution of 1-3, where each number corresponds to the following: 1 is appropriate, 2 is inappropriate and 3 is no support. To keep track of the impact that each of the scenarios make the scenarios are run separately and then compared to each other. For instance, during one simulation, which consists of all the events happening, only appropriate support is being counted and etc.
- *Perceived support update.* At each time step*,* when the random draw is 1, perceived support of the agent increases by 0.7. When the random draw is 2, perceived support of the agent decreases by 0.7. When the random draw is 3, perceived support of the agent remains constant.
- *Individual buffer.* It is determined by the perceived support level of each agent and indicates by how much perceived support buffers against distress. It is calculated with this formula: *Perceived Support update / 11.* It gives the percentage value which is multiplied by the *impact* value of the event happening at that specific time step and the product of this is subtracted from the distress level possessed by the agent at the same specific time step. It is divided by 11 to make sure that the protection is high enough when the perceived support level is high and low enough when the perceived support levels is low. At the same time, it follows the aforementioned assumption that no one is totally immune to stress and even having the maximum value of perceived support (value of 10) would not be able to eliminate all of the distress elicited by the event.
- *Low vulnerability.* If an agent belongs to the collective of low vulnerability agents and their distress level after all the events have happened is equal or higher than 40, then an agent is assigned a value of MDD onset equal to 1, meaning it is very likely that Major Depressive Disorder might have started. Otherwise, the value of MDD onset is set to 0.

Ruta Slivkaite
201805872@post.au.dk

- *High vulnerability.* The same rules apply to the collective of high vulnerability agents as to low vulnerability agents, when their distress level is equal or higher than 13.33, which ensures a threefold increased risk of MDD onset in a high vulnerability group mentioned in the literature. This would also ensure that one extremely negative event would be enough to precipitate an MDD onset when the perceived support level is very low.

Results

In this section the results of the ABM are going to be assessed by summarizing and providing descriptive statistics of the state variables of the agents after each run of simulation (4 runs in total). In order to further explore the relationships between the variables scatter plots and bar plots are going to be employed.

The results of the four received support scenarios are displayed in Table 3 below. For each vulnerability group in all received support scenarios the output of these variables were summarised: the total number of agents who were likely to experience their first MDD episode; average perceived support level with standard deviation values; average distress level with standard deviation values; total amount of received support. In the summary of the result below, the low vulnerability group will be referred to as "*L*", while the high vulnerability group as "*H*".

As expected, stressful life events lead to a considerably increased number of MDD onset in a high vulnerability group compared to a low vulnerability group in all received support scenarios confirming H1.

Appropriate received support buffered against distress the most (*Average Distress in L*= 22, sd=12; *Average Distress in H*=24, sd=12) and, thus, lowered the number of MDD onset in both low and high vulnerability groups the most *(Number of MDD onset in L*=16, *Number of MDD onset in H*=110) compared to other scenarios, while inappropriate received support buffered against distress the least (*Average Distress in L*= 30, sd=12; *Average Distress in H*=31, sd=12) and, thus, increased the number of MDD onset in both low and high vulnerability groups the most *(Number of MDD onset in L*=41, *Number of MDD onset in H*=135). These results confirmed H3. Scenarios with both kinds of

Ruta Slivkaite
201805872@post.au.dk

received support and with no support almost did not differ from each other. The visual representation of the output from the four received support scenarios is displayed in Figure 4 and Figure 5 below.

Received support scenario	*Vulnerability group*	*Number of MDD onset*	*Average perceived support*	*Average Distress*	*Total received support count*
Both appropriate (a) and inappropriate (i)	*L*	*24*	*5.1 (±3)*	*26 (±11)*	*759(a), 736(i)*
	H	*120*	*4.9 (±2.9)*	*27 (±12)*	*765(a), 739(i)*
Only appropriate	*L*	*16*	*7.8 (±2.4)*	*22 (±12)*	*754*
	H	*110*	*7.6 (±2.4)*	*24 (±12)*	*749*
Only inappropriate	*L*	*41*	*2.3(±2.3)*	*30(±12)*	*697*
	H	*135*	*1.9(±2.3)*	*31(±12)*	*765*
No received support	*L*	*31*	*5.0(±2.9)*	*26(±13)*	*0*
	H	*120*	*4.7(±3.0)*	*27(±13)*	*0*

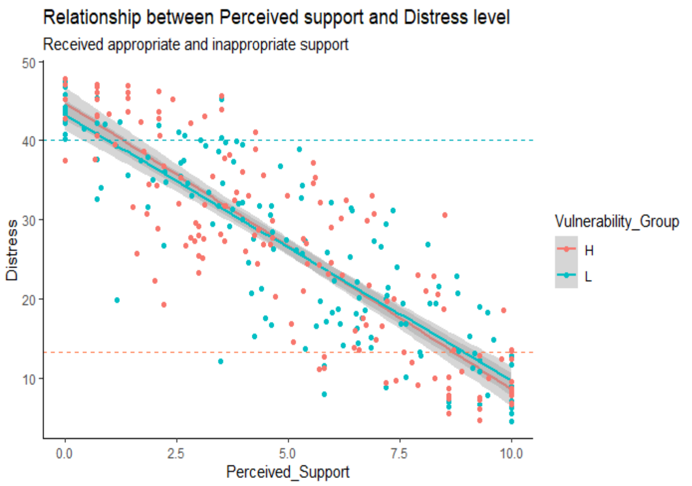
Table 3 - results of the four received support scenarios

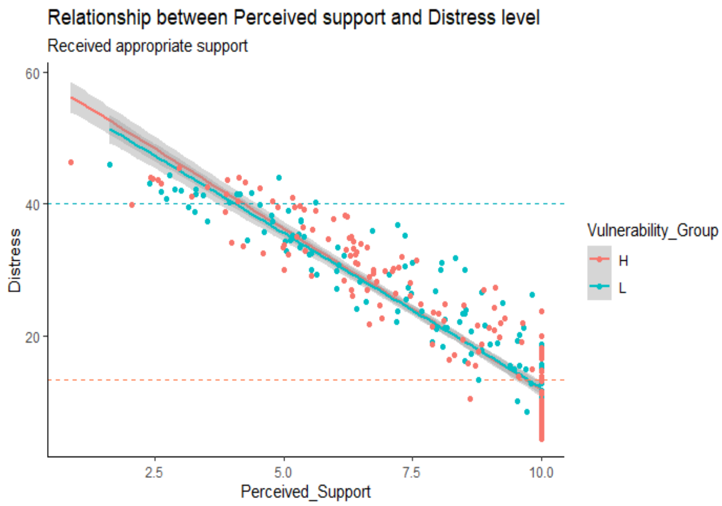
Scatter plot with the line of best fit was employed to display the relationship between perceived support level and distress level in *L* and *H* vulnerability groups. The two thresholds representing different levels of distress required for the MDD onset for H and L vulnerability groups were displayed as well (*threshold for L is 40, threshold for H is 13.33).*

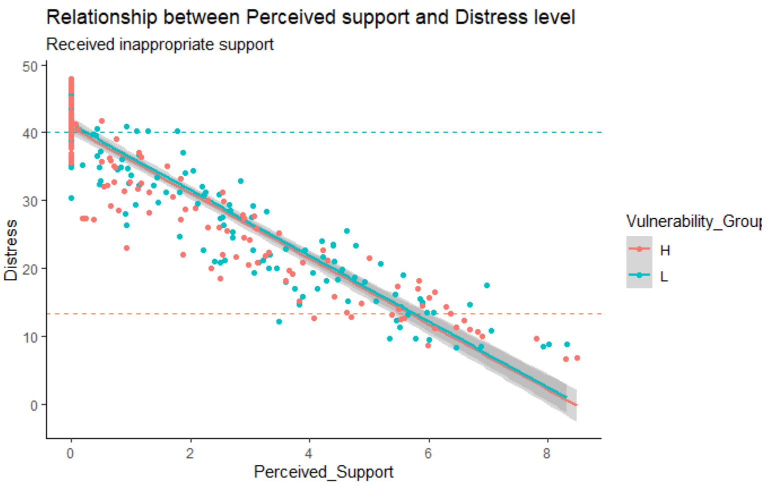
Ruta Slivkaite
201805872@post.au.dk

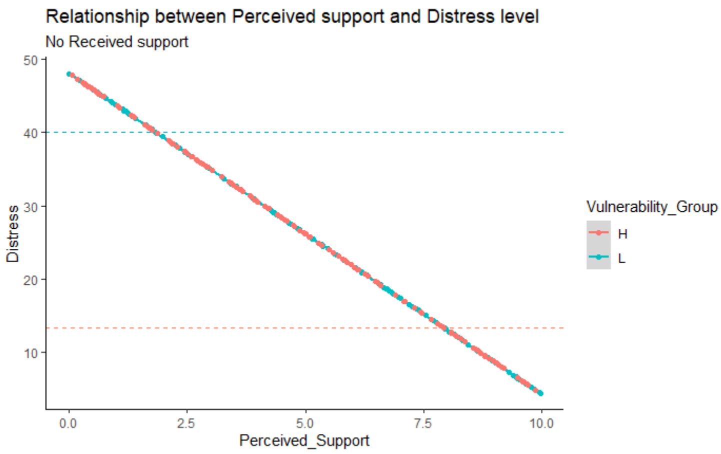
All scatter plots revealed a linear relationship between distress and perceived support level, where increased perceived support level leads to decreased distress level, which confirms H2.

We can observe that receiving appropriate support was considerably important for the high vulnerability group, because it raised perceived support level significantly to a maximum, which was important in buffering against distress and avoiding the MDD onset. In comparison, receiving inappropriate support considerably decreased perceived support level for the high vulnerability group which led to bigger distress levels and more agents experiencing MDD onset. Receiving appropriate support followed a similar trend in the lower vulnerability group, however, due to the higher threshold, less agents experienced MDD onset.









Ruta Slivkaite
201805872@post.au.dk

Figure 4 - Scatter plots displaying relationship between perceived support and distress by vulnerability group in four received support scenarios with thresholds representing different levels of distress required for the MDD onset for H and L vulnerability groups

When both kinds of support were received, we can observe more variability in perceived support and distress levels due to receiving both kinds of support, which neither extremely increased nor decreased the perceived support values. Also, lower perceived support levels led to higher distress level form high vulnerability group compared to low vulnerability group. Having received no support revealed the same linear trend for both vulnerability groups.

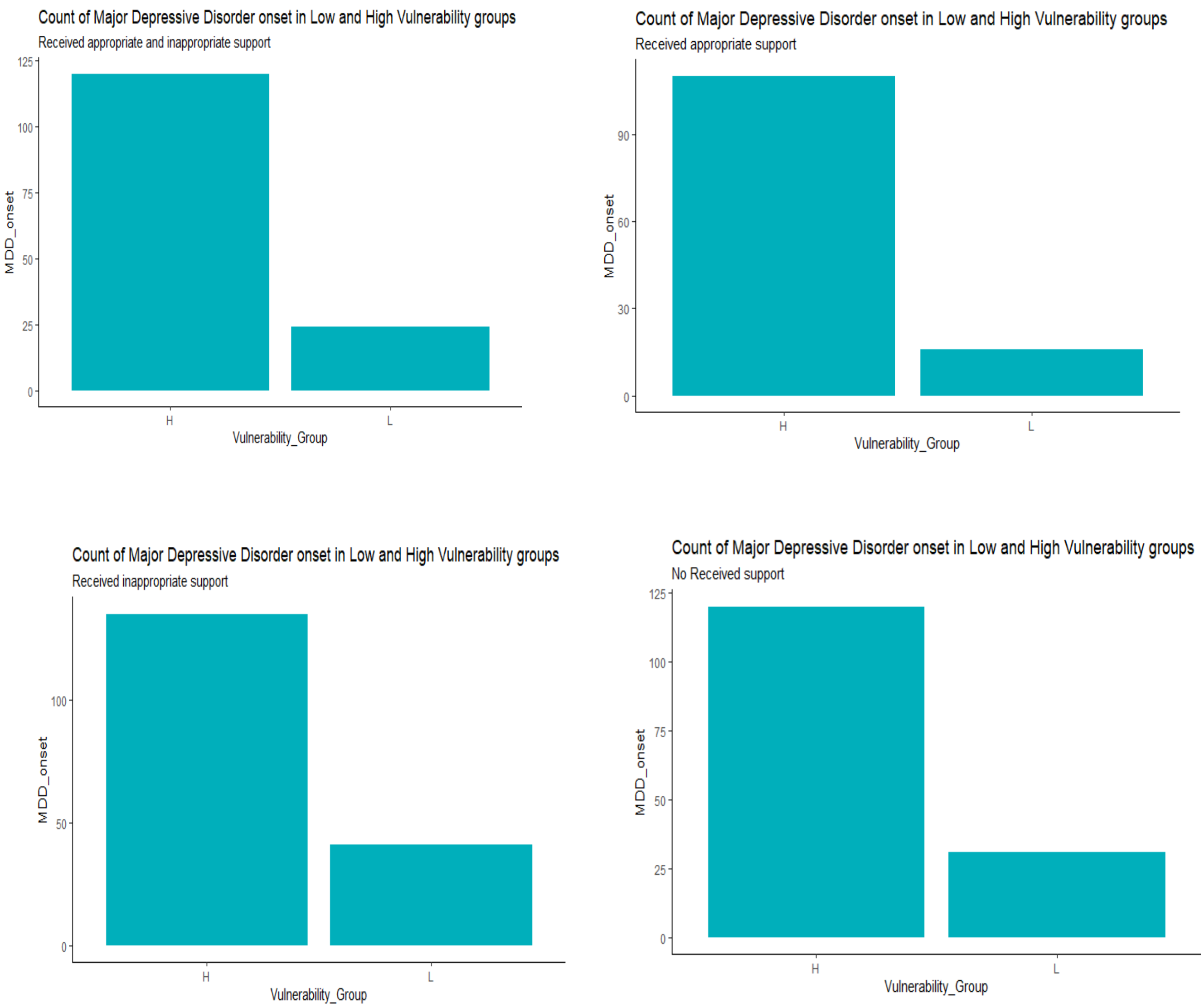


Figure 5 - Bar plots representing total number of MDD onset by vulnerability group in four received support scenarios

Ruta Slivkaite
201805872@post.au.dk

Reference list

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Hyde, J. & Mezulis, A. (2008). *The ABCs of depression: Integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression - ProQuest*. Retrieved 13 May 2020, from https://search-proquest-com.ez.statsbiblioteket.dk:12048/docview/614483601?OpenUrlRefId=info:xri/sid:primo&accountid=14468#REF_c289

Syeda, R., & Afzal, F. (2019). Perceived Social Support predicts Psychological Problems among University Students. *The International Journal of Indian Psychology*, *4*, 2349–3429.

Uchino, B. N. (2009). Understanding the Links Between Social Support and Physical Health: A Life-Span Perspective With Emphasis on the Separability of Perceived and Received Support. *Perspectives on Psychological Science*, *4*(3), 236–255. https://doi.org/10.1111/j.1745-6924.2009.01122.x

Ruta Slivkaite
201805872@post.au.dk

Appendix

Table 17.1 The seven elements of the ODD protocol (See also the ODD webpages at www.ufz.de/oesatools/odd)

ODD	ODD element	Questions to be answered
Overview	Purpose	What is the purpose of the model?
	Entities, state variables, and scales	What kind of entities are in the model? By what state variables, or attributes, are these entities characterised? What are the temporal and spatial resolutions and extents of the model?
	Process overview and scheduling	Which entities do what, in what order? When are state variables updated? How is time modelled — as discrete steps or as a continuum over which both continuous processes and discrete events can occur?
Design concepts	Design concepts	There are ten design concepts (see text). How have these concepts been taken into account in the model's design?
Details	Initialisation	What is the initial state of the model, i.e. at time $t=0$?
	Input data	What input does the model use from external sources such as data files or other models to represent processes that change over time?
	Submodels	What, in detail, are the submodels that represent the processes listed in "Process overview and scheduling"? What are the model parameters, their dimensions, and reference values? How were submodels designed or chosen, tested, and parameterised?

Figure 2 - ODD protocol with questions (Grimmet al., 2010)

	Elements of the ODD protocol
Overview	1. Purpose
	2. Entities, state variables, and scales
	3. Process overview and scheduling
Design concepts	4. Design concepts - Basic principles - Emergence - Adaptation - Objectives - Learning - Prediction - Sensing - Interaction - Stochasticity - Collectives - Observation
Details	5. Initialization
	6. Input data
	7. Submodels

Figure 3 - ODD protocol (Grimmet al., 2010)