Computational Modeling - Week 3 - Assignment 2 - Part 1

Signe Kløve and Thea Rolskov Sloth

Link to GitHub:
https://github.com/thearol/ComMod/blob/master/2_bayesian_estimates/Thea_Assignment2_Part1.Rmd

In this assignment we learn how to assess rates from a binomial distribution, using the case of assessing your teachers' knowledge of CogSci

First part

You want to assess your teachers' knowledge of cognitive science. "These guys are a bunch of drama(turgist) queens, mindless philosophers, chattering communication people and Russian spies. Do they really know CogSci?", you think.

To keep things simple (your teachers should not be faced with too complicated things): - You created a pool of equally challenging questions on CogSci - Each question can be answered correctly or not (we don't allow partially correct answers, to make our life simpler). - Knowledge of CogSci can be measured on a scale from 0 (negative knowledge, all answers wrong) through 0.5 (random chance) to 1 (awesome CogSci superpowers)

This is the data: - Riccardo: 3 correct answers out of 6 questions - Kristian: 2 correct answers out of 2 questions (then he gets bored) - Josh: 160 correct answers out of 198 questions (Josh never gets bored) - Mikkel: 66 correct answers out of 132 questions

Questions:

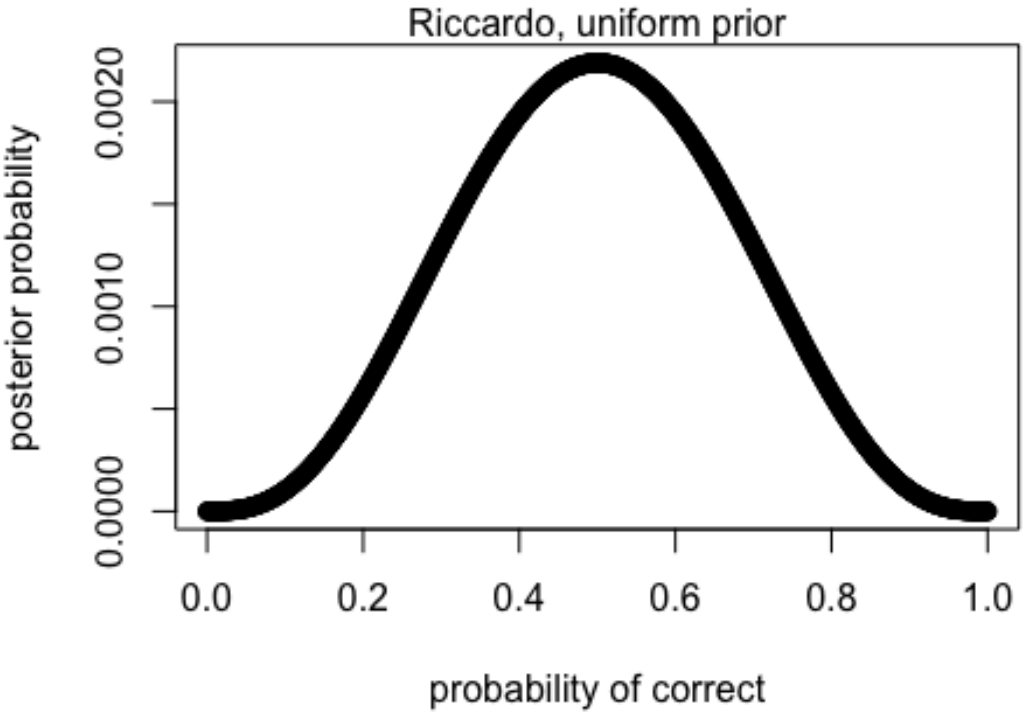
1. What's Riccardo's estimated knowledge of CogSci? What is the probability he knows more than chance (0.5) [try figuring this out. if you can't peek into chapters 3.1 and 3.2 and/or the slides]?

```
sum( samples_ric > 0.5 ) / length (samples_ric)
```

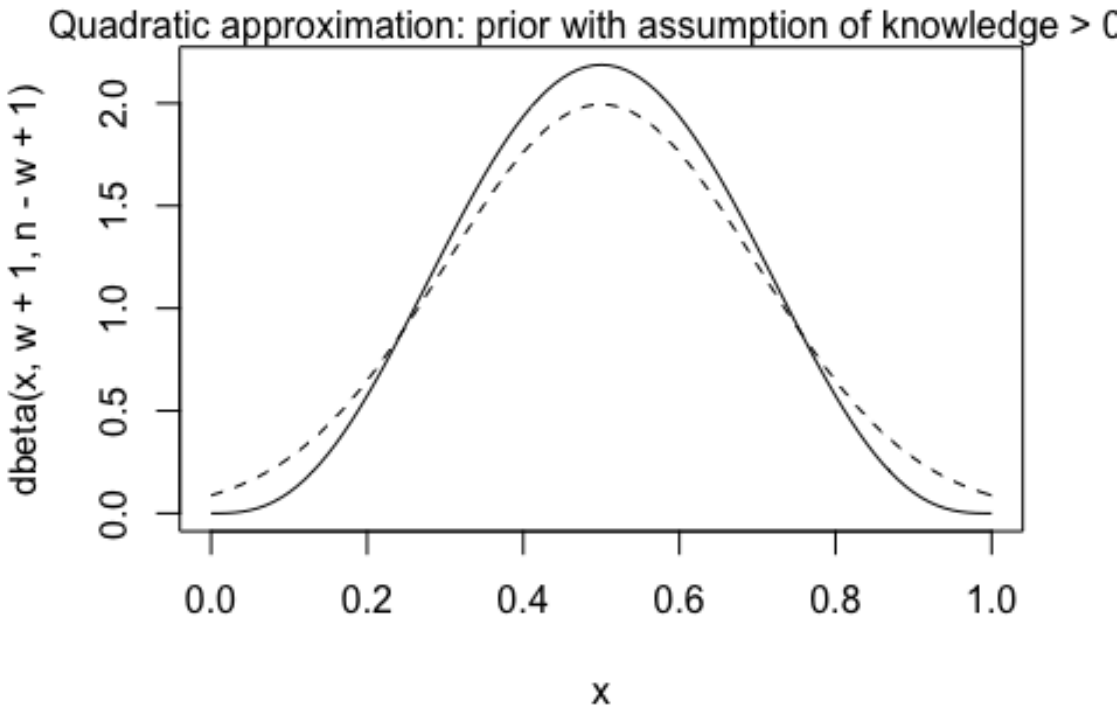
```
## [1] 0.4905
```

The propbability of Riccardo knowing more than chance depends on the prior. If it is the uniform prior then there is a 50% chance that he knows more than 0.5. If we use a prior which assigns 0 to grid-values below 0.5 and 1 to grid-values above 0.5, then there is a 100% chance that he performs above chance.

- First implement a grid approximation (hint check paragraph 2.4.1!) with a uniform prior, calculate the posterior and plot the results



- Then implement a quadratic approximation (hint check paragraph 2.4.2!).

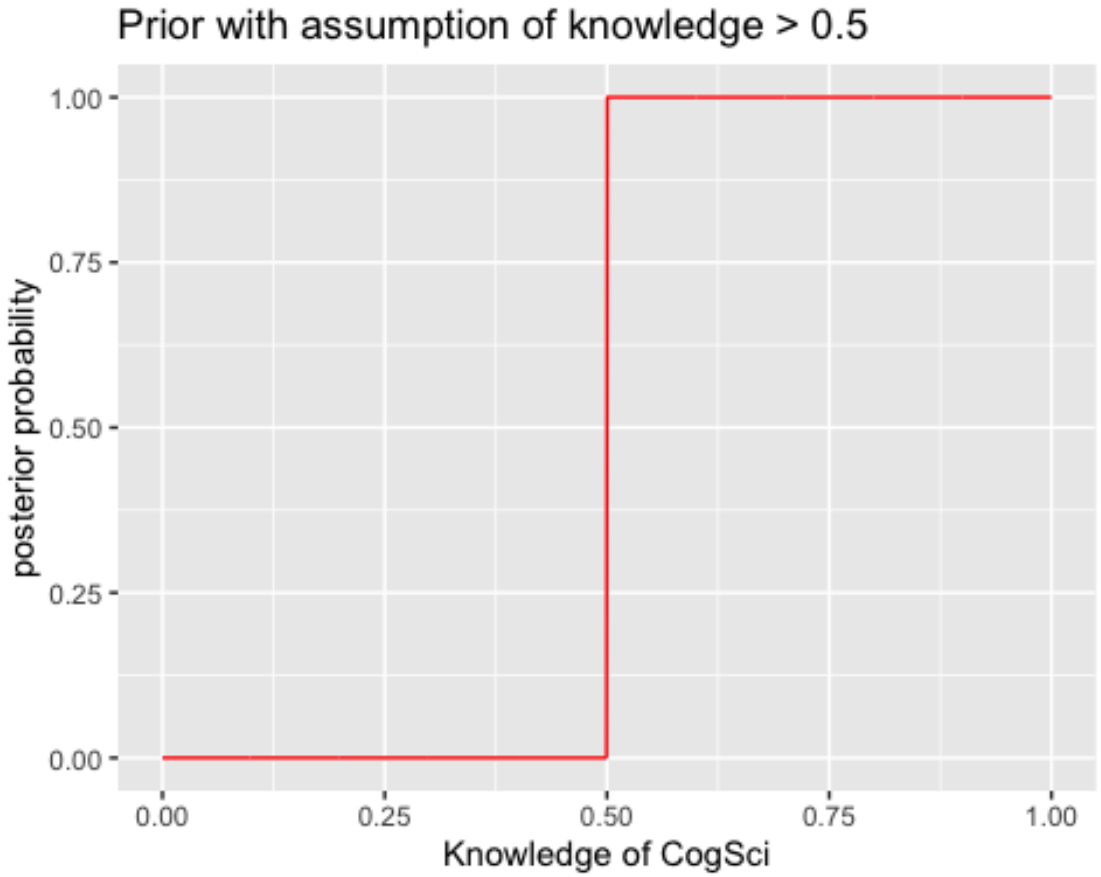


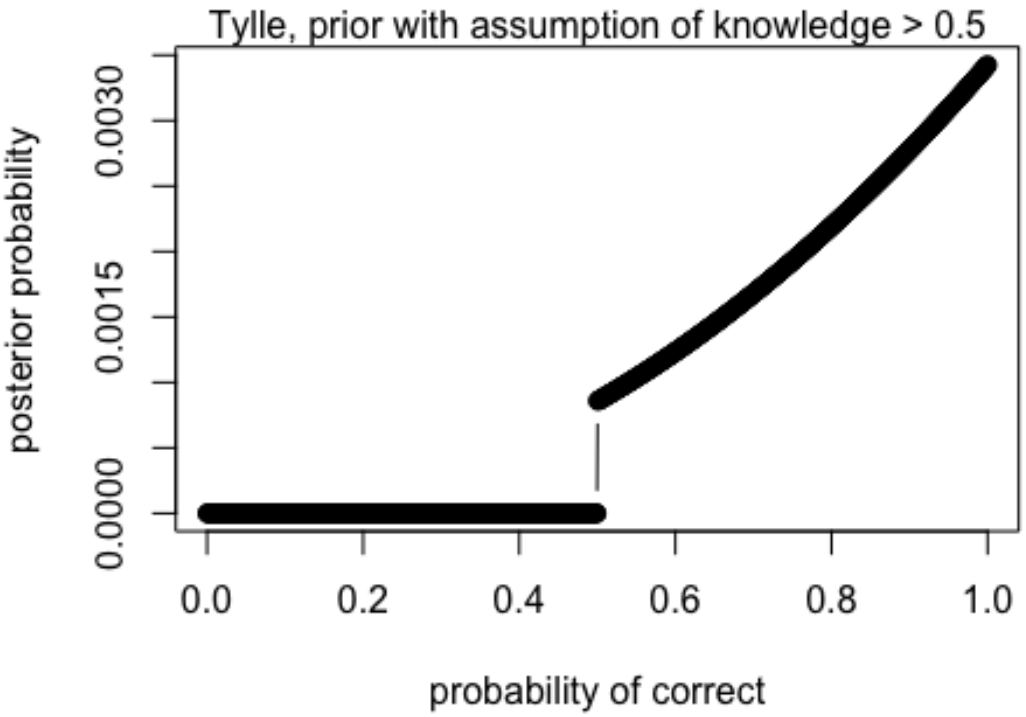
N.B. for the rest of the exercise just keep using the grid approximation (we'll move to quadratic approximations in two classes)

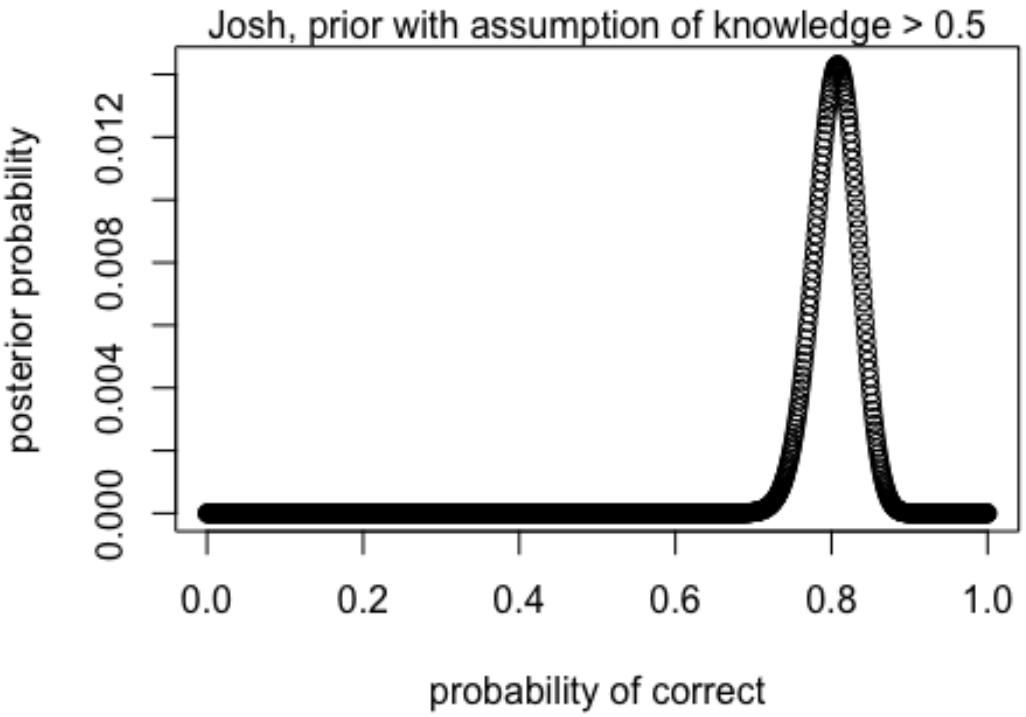
2. Estimate all the teachers' knowledge of CogSci. Who's best? Use grid approximation. Comment on the posteriors of Riccardo and Mikkel.

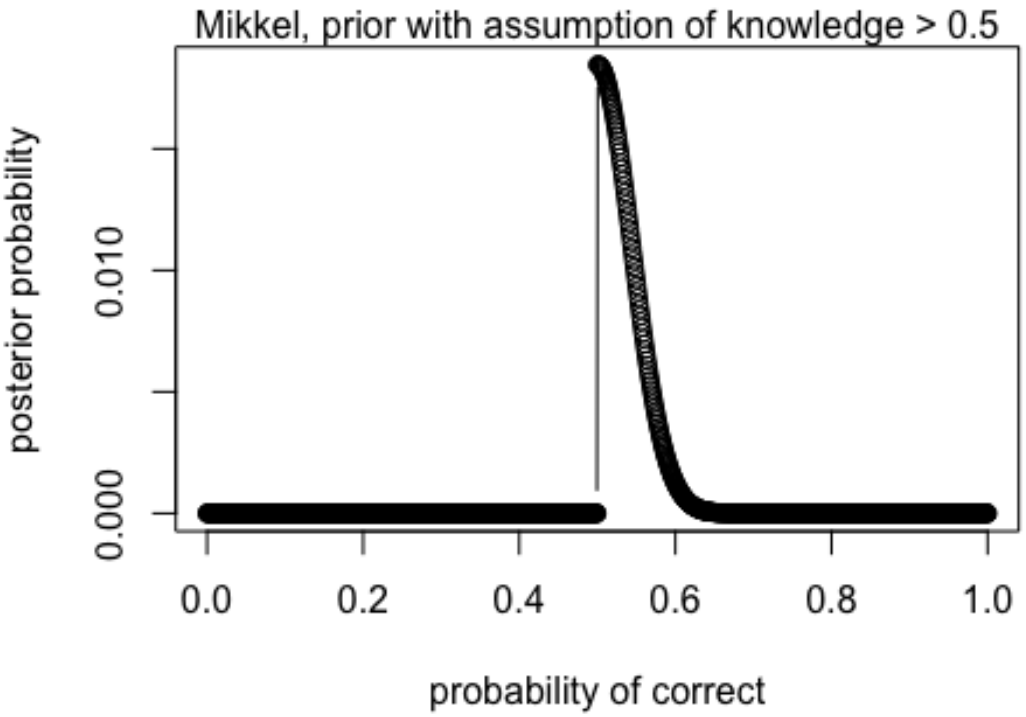
2a. Produce plots of the prior, and posterior for each teacher.

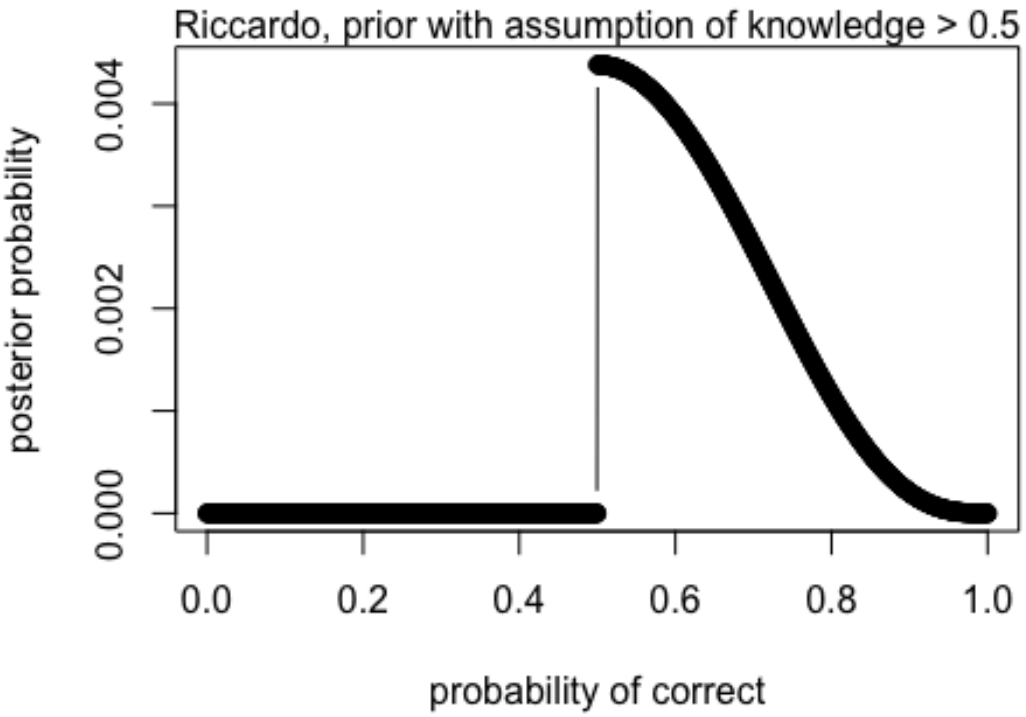
From now on we will use the following prior that assumes the teachers will perform better than chance (which is basically the same as having zero CogSci knowledge)


Prior with assumption of knowledge > 0.5
posterior probability
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
Knowledge of CogSci


Tylle, prior with assumption of knowledge > 0.5
posterior probability
0.0000
0.0015
0.0030
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct


Josh, prior with assumption of knowledge > 0.5
posterior probability
0.000
0.004
0.008
0.012
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct


Mikkel, prior with assumption of knowledge > 0.5
posterior probability
0.000
0.010
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct



In percentages Riccardo and Mikkel answer roughly equally bad. However, since we have many more datapoints from Mikkel, as the posterior distribution shows, we can be much more certain in assessing Mikkel's CogSci knowledge. This is shown in the area under the curve being more narrow compared to Riccardo's.

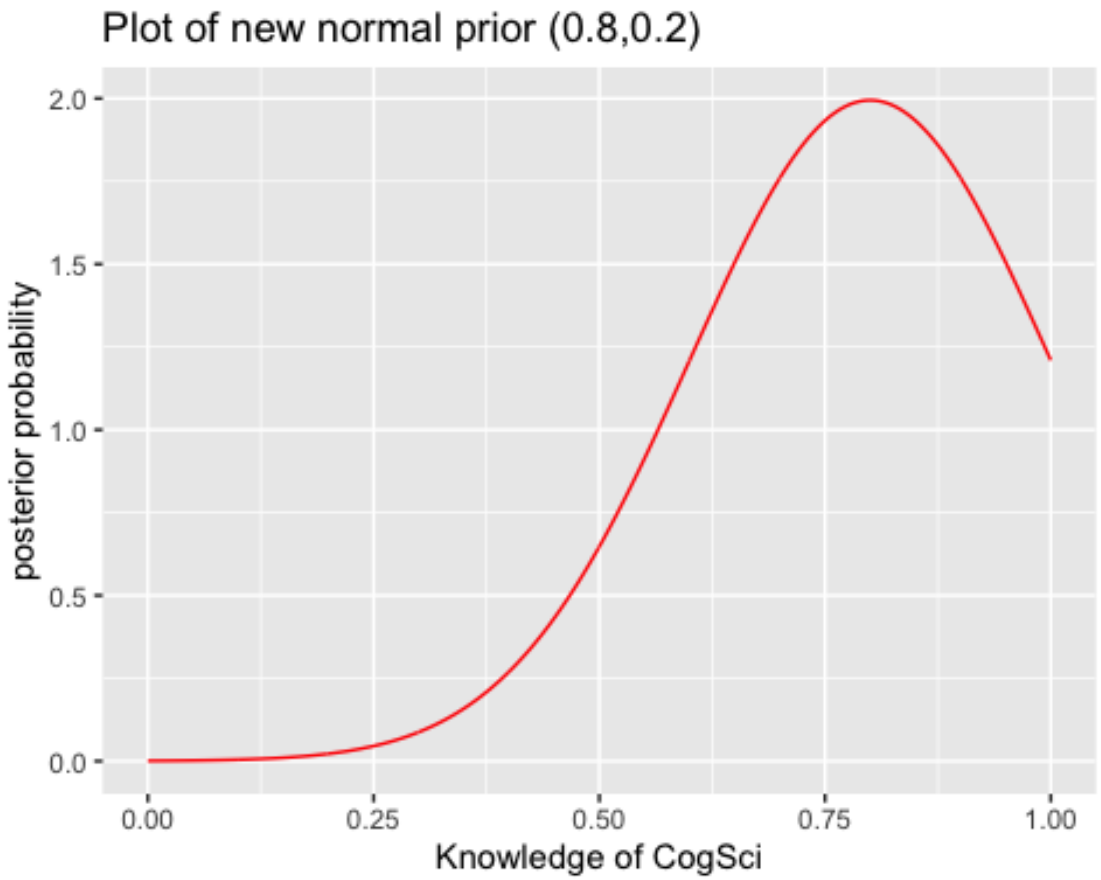
Who's best?

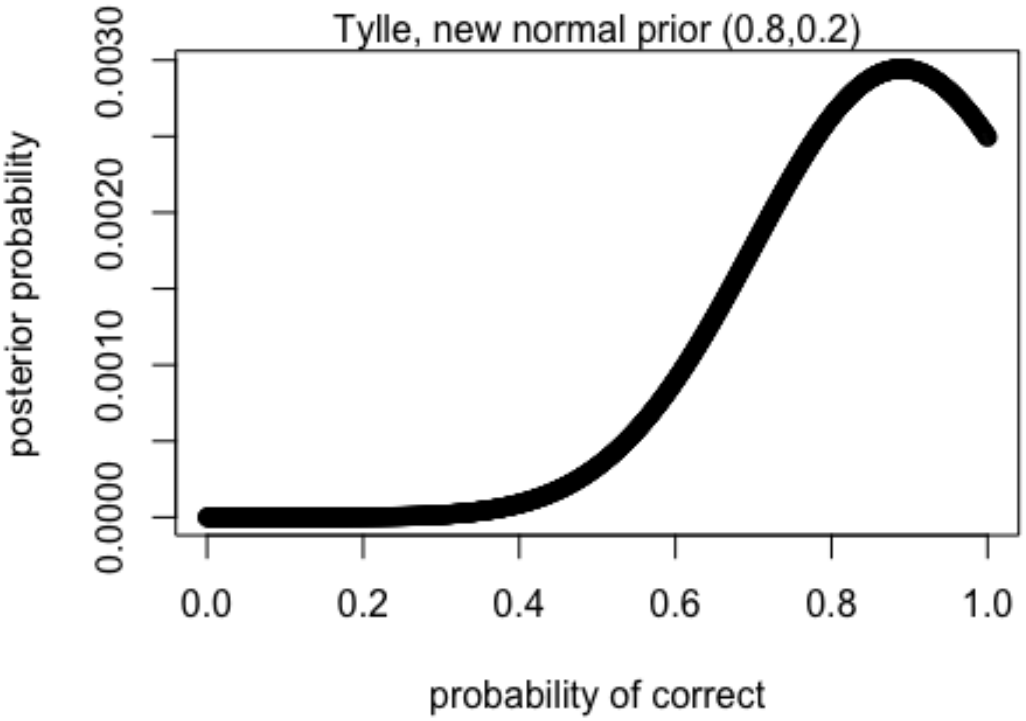
Assessing the posterior distribution, we are quite certrain that Josh's knowledge lays around 0.8, which is somewhat impressive. However, looking at Tylle's curve we are very uncertain - it could be 0.5, but also 1 (awesome). This is due to the fact that we have sucj few datapoints from Tylle, resulting in high uncertainty.

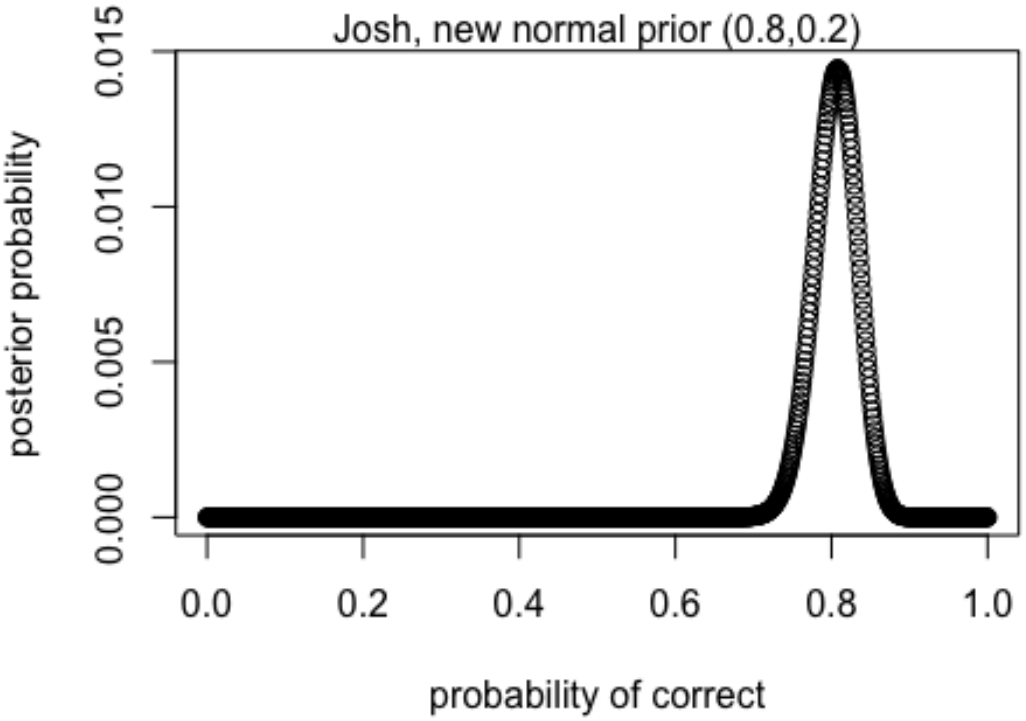
3. Change the prior. Given your teachers have all CogSci jobs, you should start with a higher appreciation of their knowledge: the prior is a normal distribution with a mean of 0.8 and a standard deviation of 0.2. Do the results change (and if so how)?

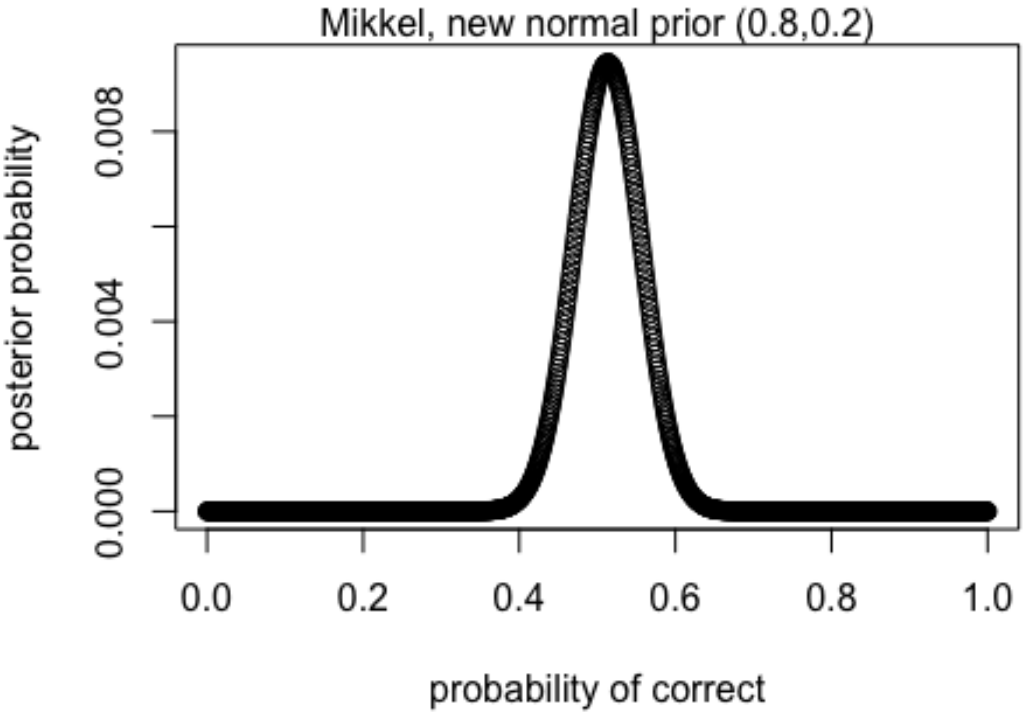
3a. Produce plots of the prior and posterior for each teacher.

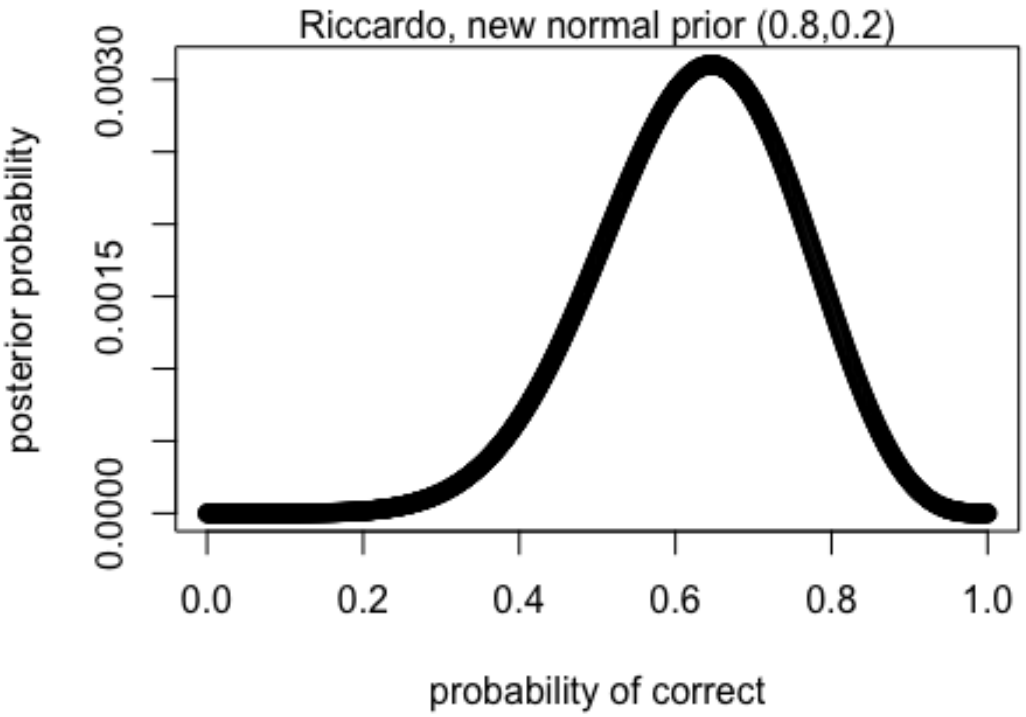
```
new_prior <- dnorm(grid,0.8,0.2)
# plot(new_prior)
#plot of new prior
```




Tylle, new normal prior (0.8,0.2)
posterior probability
0.0000
0.0010
0.0020
0.0030
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct


Josh, new normal prior (0.8,0.2)
posterior probability
0.000
0.005
0.010
0.015
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct


Mikkel, new normal prior (0.8,0.2)
posterior probability
0.000
0.004
0.008
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct



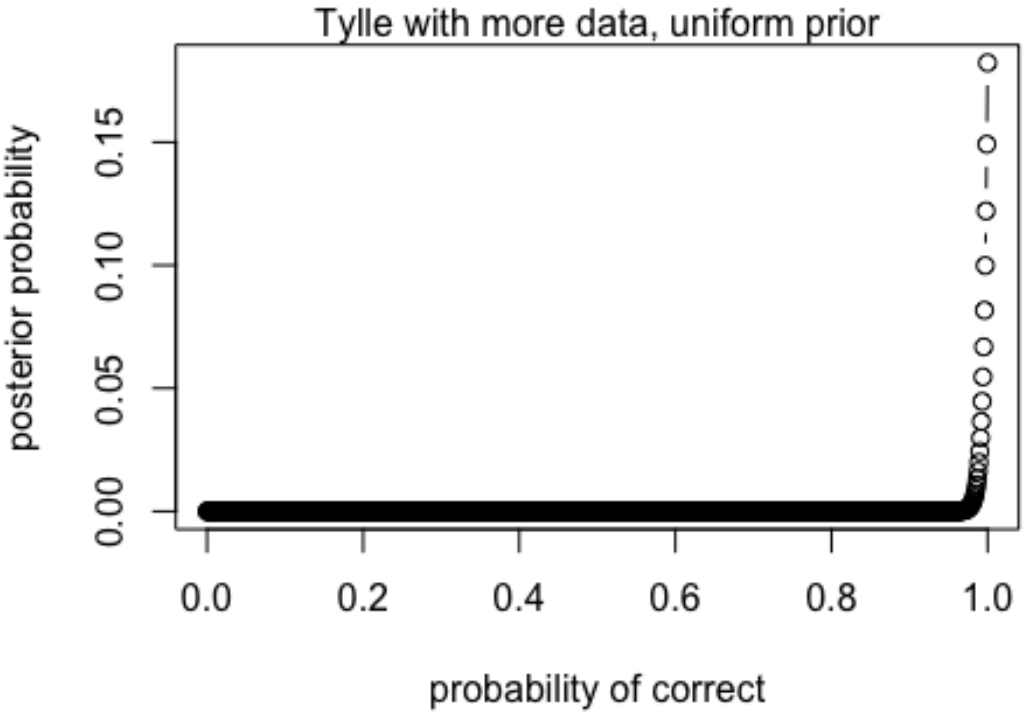
Change in results: Yes, they change. At least for Tylle and Riccardo. We can see that the prior affects the posterior more, when there are few datapoints. Josh and Mikkel remain more or less the samme, since there are many datapoints, and thereby are not so manipulative by the prior.

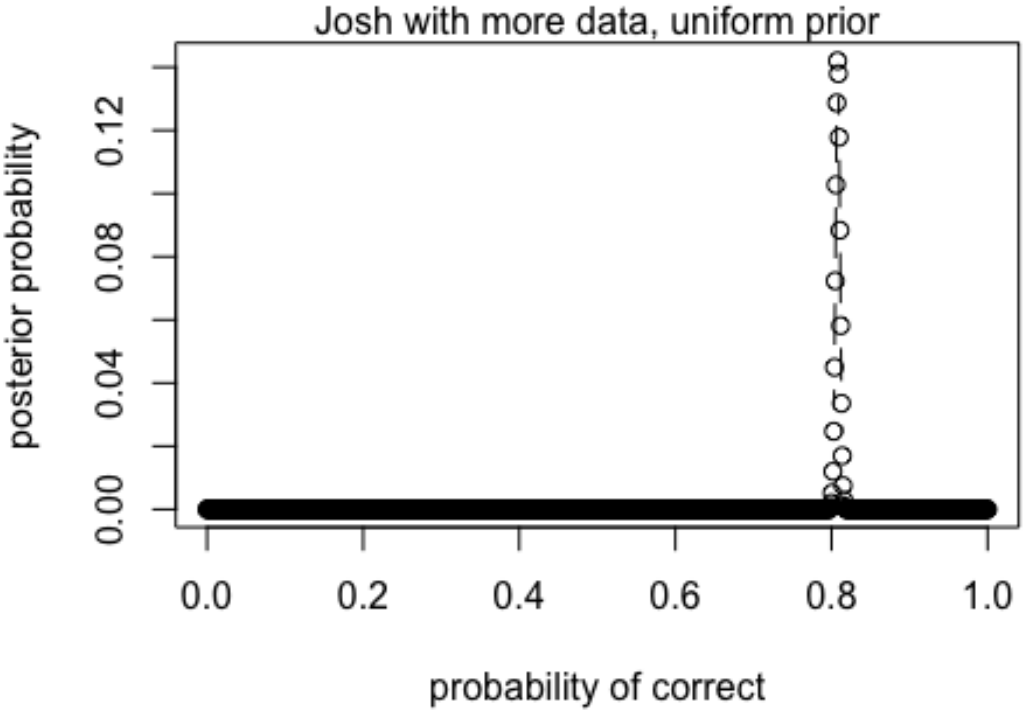
4. You go back to your teachers and collect more data (multiply the previous numbers by 100). Calculate their knowledge with both a uniform prior and a normal prior with a mean of 0.8 and a standard deviation of 0.2. Do you still see a difference between the results? Why?

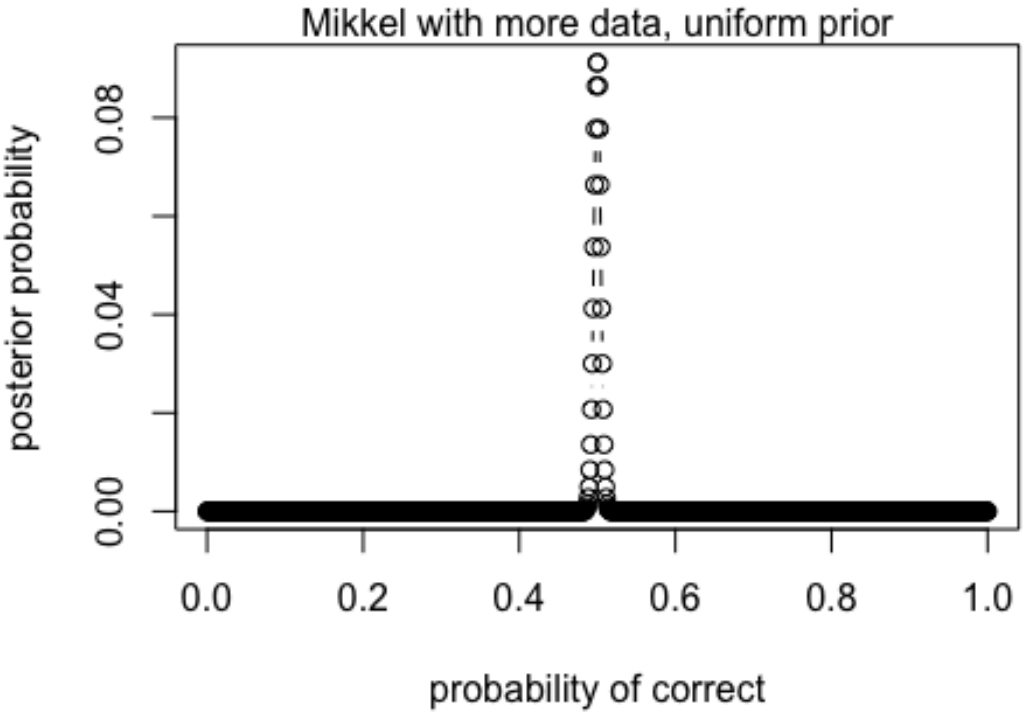
```
#multiply all data with 100
new_d <- d*100
```

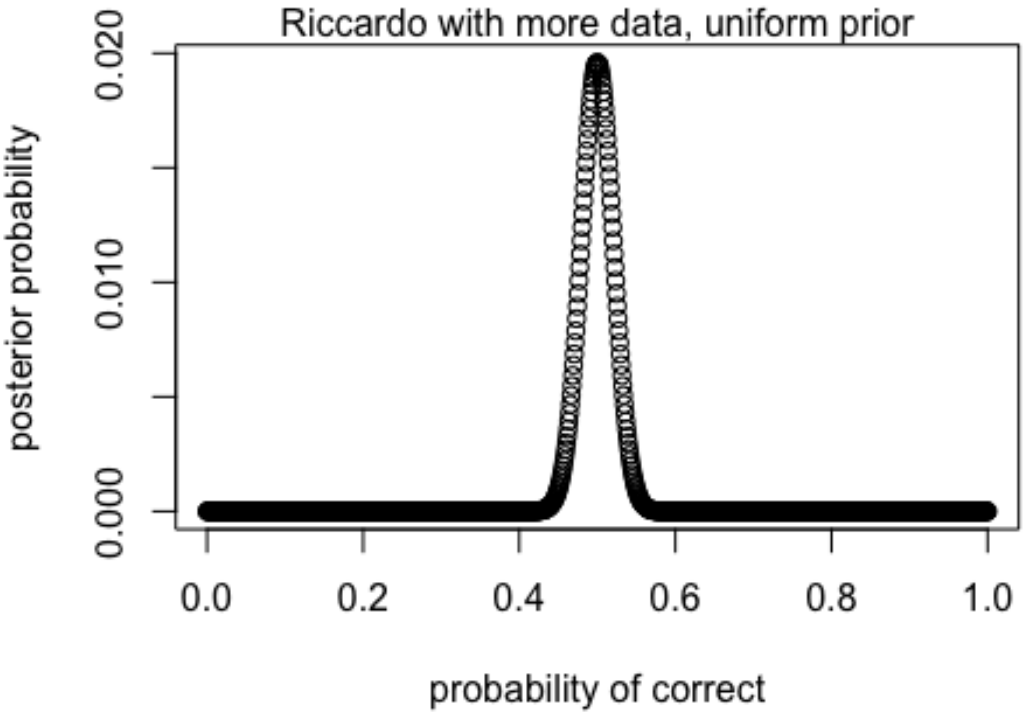
```
##   Correct Questions Teacher
## 1     300       600      RF
## 2     200       200      KT
## 3   16000     19800      JS
## 4    6600     13200      MW
```

Uniform prior

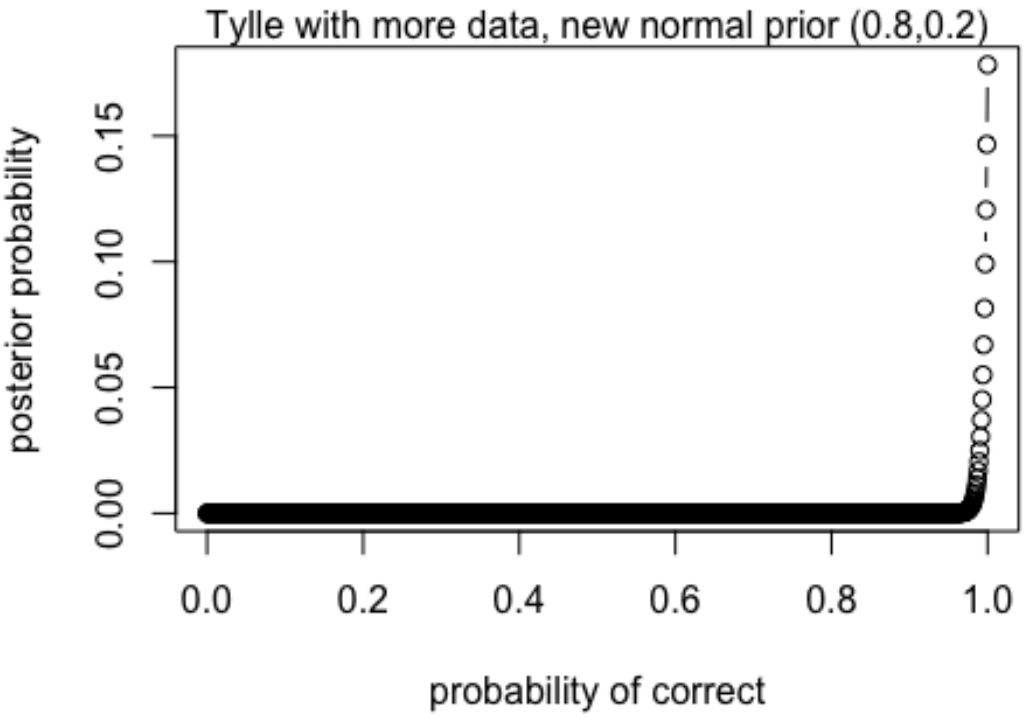


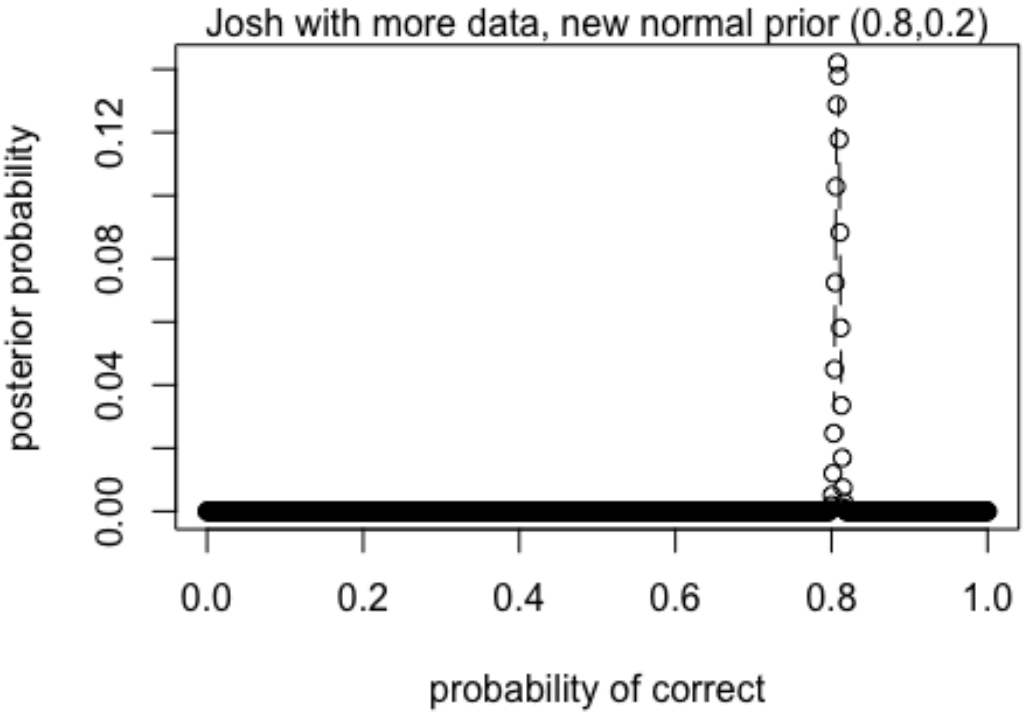

Josh with more data, uniform prior
posterior probability
0.00
0.04
0.08
0.12
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct

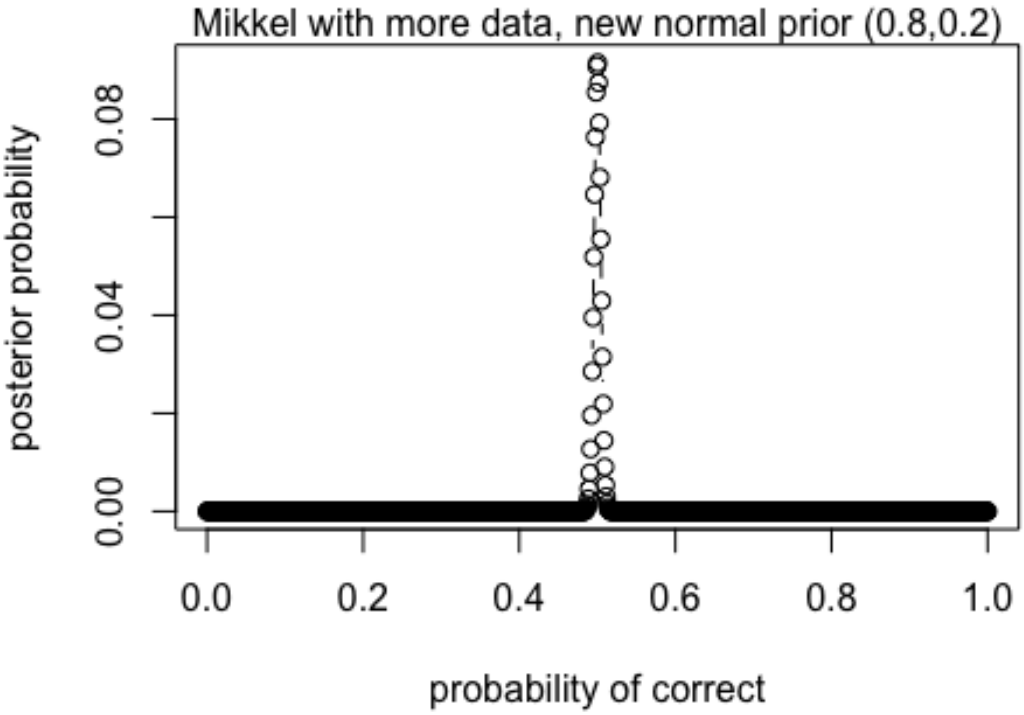

Mikkel with more data, uniform prior
posterior probability
0.00
0.04
0.08
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct

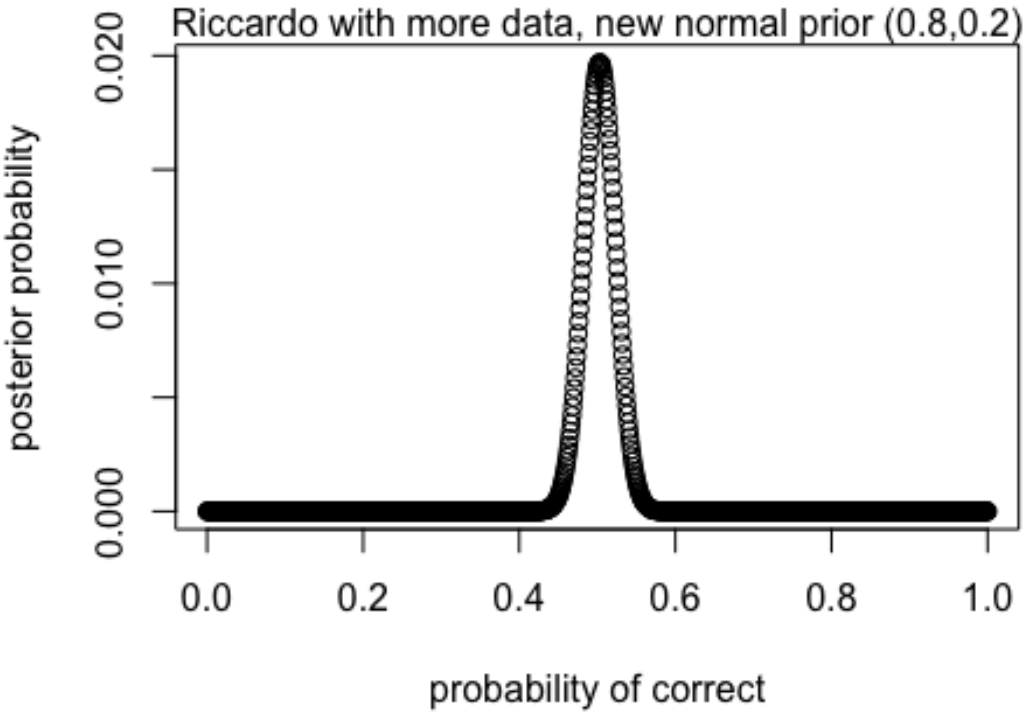

Riccardo with more data, uniform prior
posterior probability
0.000
0.010
0.020
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct

Normal prior:




Josh with more data, new normal prior (0.8,0.2)
posterior probability
0.00
0.04
0.08
0.12
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct


Mikkel with more data, new normal prior (0.8,0.2)
posterior probability
0.00
0.04
0.08
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct

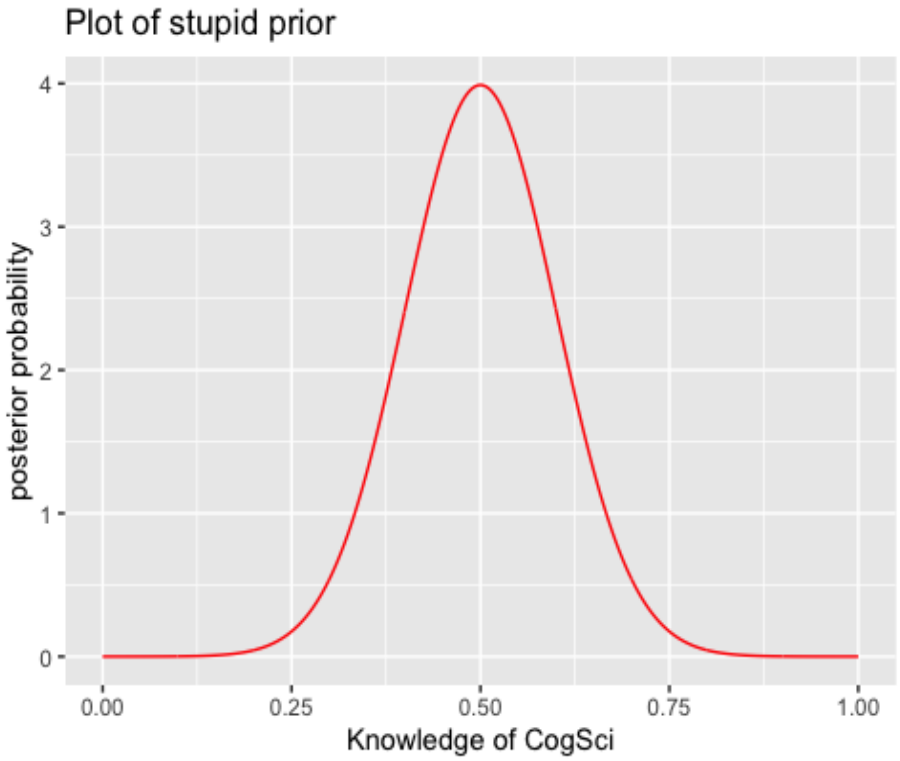


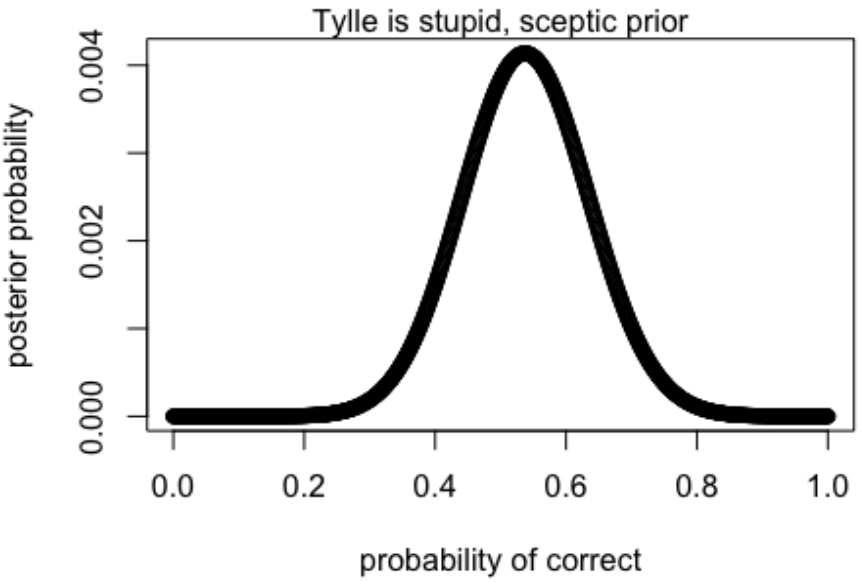
The impact of the prior becomes very small when we have so much data.

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?

Make a prior of a normal distribution that peaks at 0.5 and with a quite small SE, since we are pretty sure that they are useless.

```
stupid_prior <- dnorm(grid,0.5,0.1)
```


Plot of stupid prior
posterior probability
0
1
2
3
4
0.00
0.25
0.50
0.75
1.00
Knowledge of CogSci


Tylle is stupid, sceptic prior
posterior probability
0.000
0.002
0.004
0.0
0.2
0.4
0.6
0.8
1.0
probability of correct

It works!