The script to allow you to simulate repeated *t* tests yourself.

First download R and install it (from http://www.r-project.org/)

Then download R Studio (from http://www.rstudio.com/),

Create a directory/folder named something convenient like R-progs.

Dorothy Bishop has provided download instructions, and a short introduction to R, at http://www.slideshare.net/deevybishop/learning-r-while-exploring-statistics

Put the script, *two_sample-simulation.R*, into that directory, start R Studio and open the script.

R Studio Notes

Lines that begin with # are comments, not code.

All variable names start with my (a curious convention in R that's intended to avoid clashes with pre-defined variables) the rest of the name is more informative.

All you have to do is to set the inputs, which are gathered together between

#START INPUTS

and #END OF INPUTS

Type in the number of simulations that you want to do (overwrite the value of *mynsim*=100000 if you want a different number)

The *t* test compares the means of two groups of normally-distributed variables. The random values are simulated separately for each *t* test. There are *myn* observations in each group (*myn*=16 in the script, which gives a power of 0.78 for the tests). Put in the value you want.

Each sample is generated randomly such that the observations are normally distributed with mean *mymu1* and standard deviation *mysd1* for group 1 and

mean *mymu2* and standard deviation *mysd2* for group 2.

In the example script the standard deviations are 1 for both groups, *mymu1* = 0 and

mymu2 =1, so these numbers specify data with means that differ by one standard deviation as shown in Fig 4 a, b in the paper. If you want to see what happens when the null hypothesis is true, then set *mymu1* and *mymu2* to be equal (e.g. both zero).

Set *myPmin* and *myPmax.* For a conventional test in which P ≤ 0.05 is taken as "significant", set *myPmin* = 0 and *myPmax* = 0.05. To look only at tests for which *P* is close to 0.05, set *myPmin* = 0.045 and *myPmax* = 0.05, as described in the paper.

Finally specify the name of the text file that records the output. This is done in the line

```
outfile="run1.txt"              #name for output file.
```

Change the name in inverted commas to anything you want.

Now you are ready to run the script. Hit ctrl-alt-R and the script will run. It takes about 3.5 minutes for 100,000 simulations on my Windows 7 laptop and 2.2 minutes on Windows 7 desktop.
Alternatively highlight the whole script (ctrl-A) and click the run button (the run button runs the highlighted sections of the script).

Four graphs appear, each in a separate window. Before the simulations, graphs of the distributions (of observations and of means) which are sampled, as in Fig 4 a,b appear. After the simulations graphs appear that show the distribution of the 100,000 differences between means, and the distribution of the 100,000 *P* values, from each of the *mynsim* 'experiments'.

Any graph that you want to keep must be saved manually.

The output file (run1.txt in the example) appears in the same directory in which the script resides. Here is an example of the text file (the numbers will differ slightly when you repeat the test, because the data are random variables, but if you do a large enough number of simulations, the variability should small).
Here is an example of run1.txt

INPUTS
number of simulation = 1e+05
number obs per sample = 16
true mean for sample 1 = 0
true mean for sample 2 = 1
true SD (same for both samples) = 1
Sig if P is between Pmin and Pmax = 0 to 0.05

OUTPUTS
power = 0.7813965 for P = 0.05

Number of 'sig' differences (P between Pmin-Pmax) = 78244 = 78.244 percent

Number of P <= 0.001 = 23565 = 23.565 percent
Number of 0.001 < P <= 0.01 = 30404 = 30.404 percent
Number of 0.01 < P <= 0.05 = 24275 = 24.275 percent
Number of P > 0.05 = 21756 = 21.756 percent
Number of P <= 0.05 = 78244 = 78.244 percent

Obs diff between means for 'sig' results = 1.128788 True value = 1

Number of times sig is followed by non-sig (P=0.05) = 17061 = 21.80487 percent
Number of times sig is followed by sig (P=0.05) = 61183 = 78.19513 percent
Number of times sig is followed by non-sig (P=0.01) = 24846 = 46.03754 percent
Number of times sig is followed by sig (P=0.01) = 29123 = 53.96246 percent