8TH IEA INTERNATIONAL RESEARCH CONFERENCE

The Differences that Make the Difference

26-28 of June 2019 with a pre-conference on secondary data analysis 24-25 of June

DANISH SCHOOL OF EDUCATION
AARHUS UNIVERSITY

10 OCTOBER 2018

CHRISTIAN CHRISTRUP KJELDSEN
ASSOCIATE PROFESSOR

IRC 2019 AARHUS UNIVERSITY, DPU, COPENHAGEN

DIFFERENCES THAT MAKE A DIFFERENCE?

A short teaser of my approach – others will be made I am sure

I will argue for revisiting an old, but often forgotten discussion on the difference between statistical significance and substantial significance.

A space with three continuums

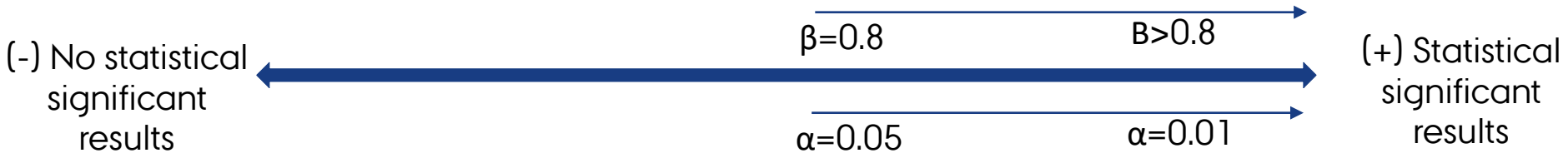
STATISTICAL SIGNIFICANCE, POWER AND EFFECT SIZES

The development of statistical test theory in the early 1900s and the new critique of statistical significance.

“From time to time we must remind ourselves that statistical significance refers only to the reliability of an obtained result, the confidence with which a null hypothesis may be rejected or the probability that a Type I error has been committed.” (Levy, 1967 p.37)

> Now, it should be noted that in the controversy over the use of tests of significance, there is no disagreement that the confusion of statistical significance with substantive importance is a sin (Gold D. 1969, p.44)

FIRST CONTINUUM WITHIN QUANTITATIVE EDUCATIONAL RESEARCH



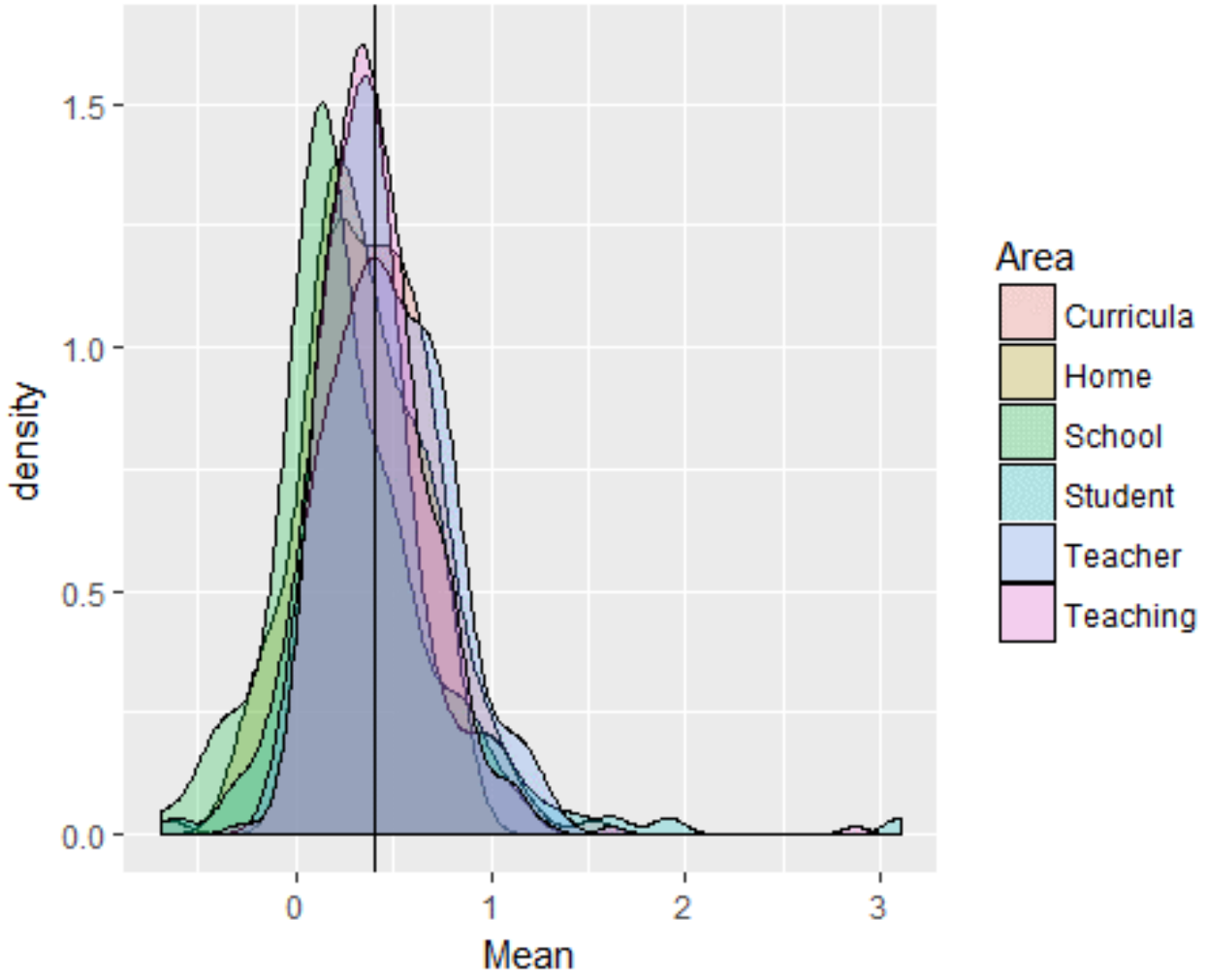
SECOND CONTINUUM WITHIN QUANTITATIVE EDUCATIONAL RESEARCH (THE OUTCOMES)

(-) Substantial significance

(+) Substantial significance: (IEA improving education – improvement of what?)

JOHN HATTIE´S META ANALYSIS

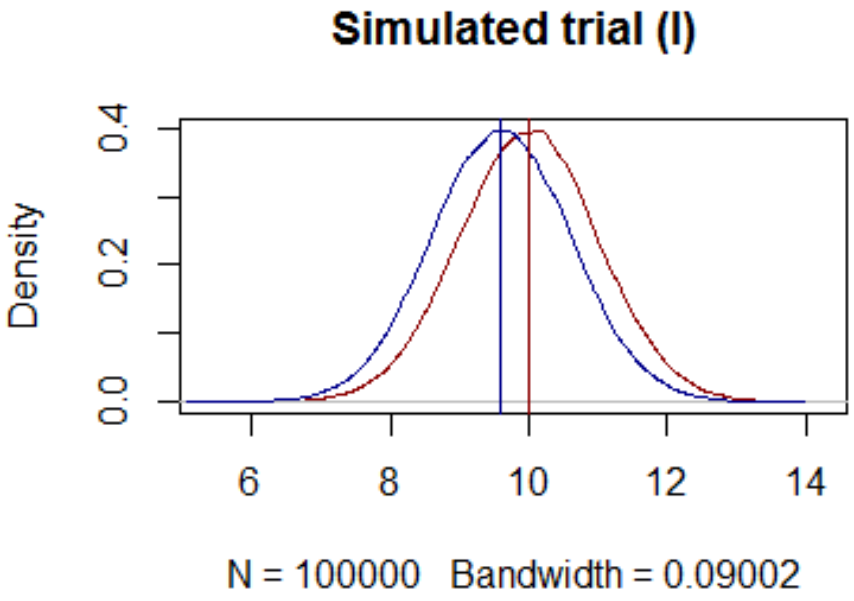
"Innovations are more than teaching: Teachers average an effect of d = 0.20 to d = 0.40 per year on student achievement. This h-point of d = 0.40 does not mean that this is the typical effect of teaching or teachers. It does not mean that merely placing a teacher in front of a class would lead to an improvement of 0.40 standard deviations." (Hattie 2009, s. 17)

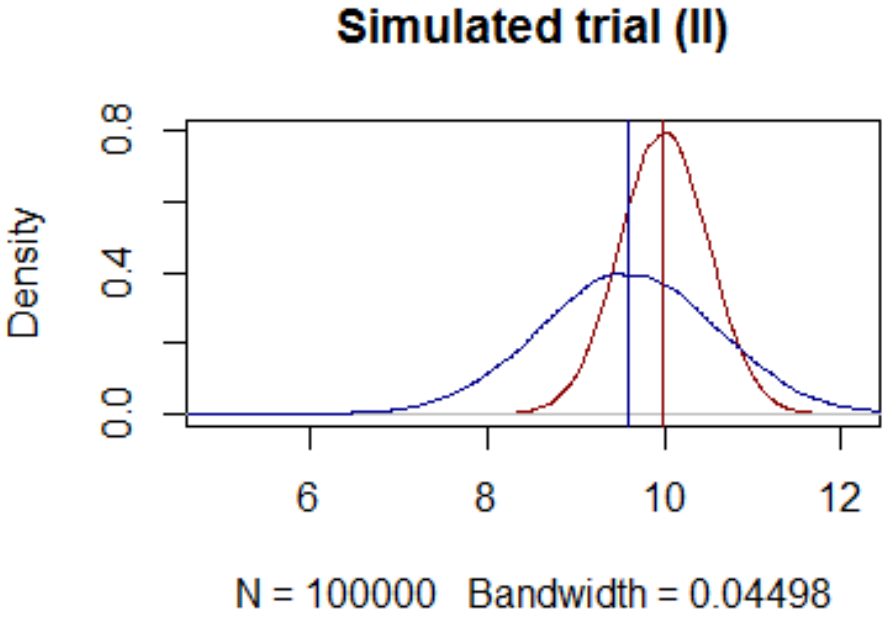


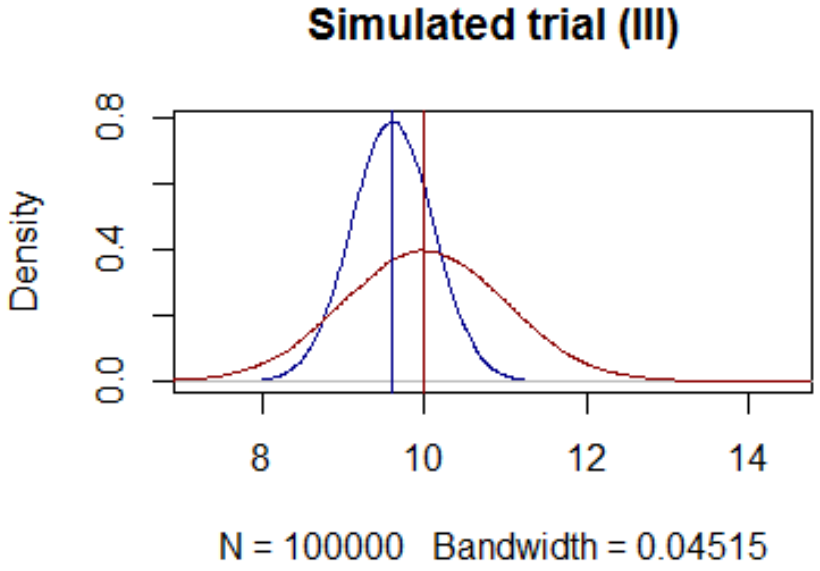
Effect Size alone: a moral philosophical problem of social justice in education?

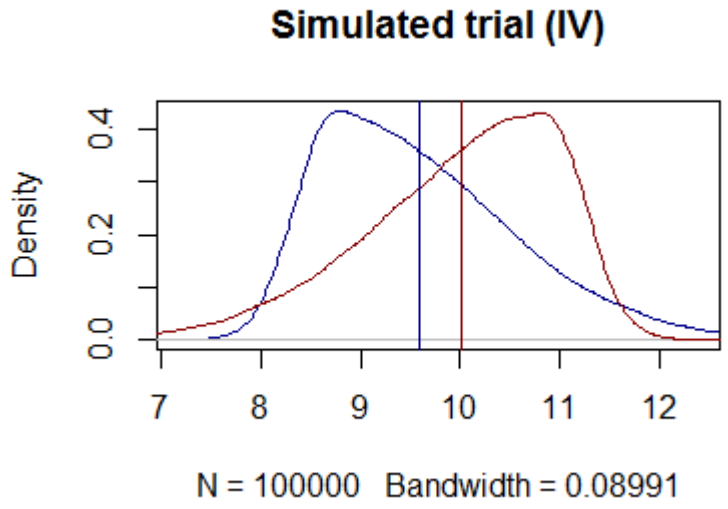
Four **different** scenarios, but with **similar** Effect Sizes – a difference that makes a difference?

CHRISTIAN CHRISTRUP KJELDSEN
ASSOCIATE PROFESSOR

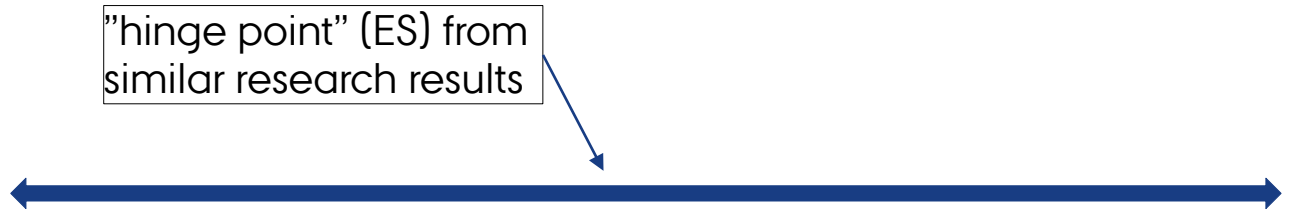

Simulated trial (I)
Density
0.0
0.2
0.4
6
8
10
12
14
N = 100000 Bandwidth = 0.09002


Simulated trial (II)
Density
0.0
0.4
0.8
6
8
10
12
N = 100000 Bandwidth = 0.04498


Simulated trial (III)
Density
0.0
0.4
0.8
8
10
12
14
N = 100000 Bandwidth = 0.04515


Simulated trial (IV)
Density
0.0
0.2
0.4
7
8
9
10
11
12
N = 100000 Bandwidth = 0.08991

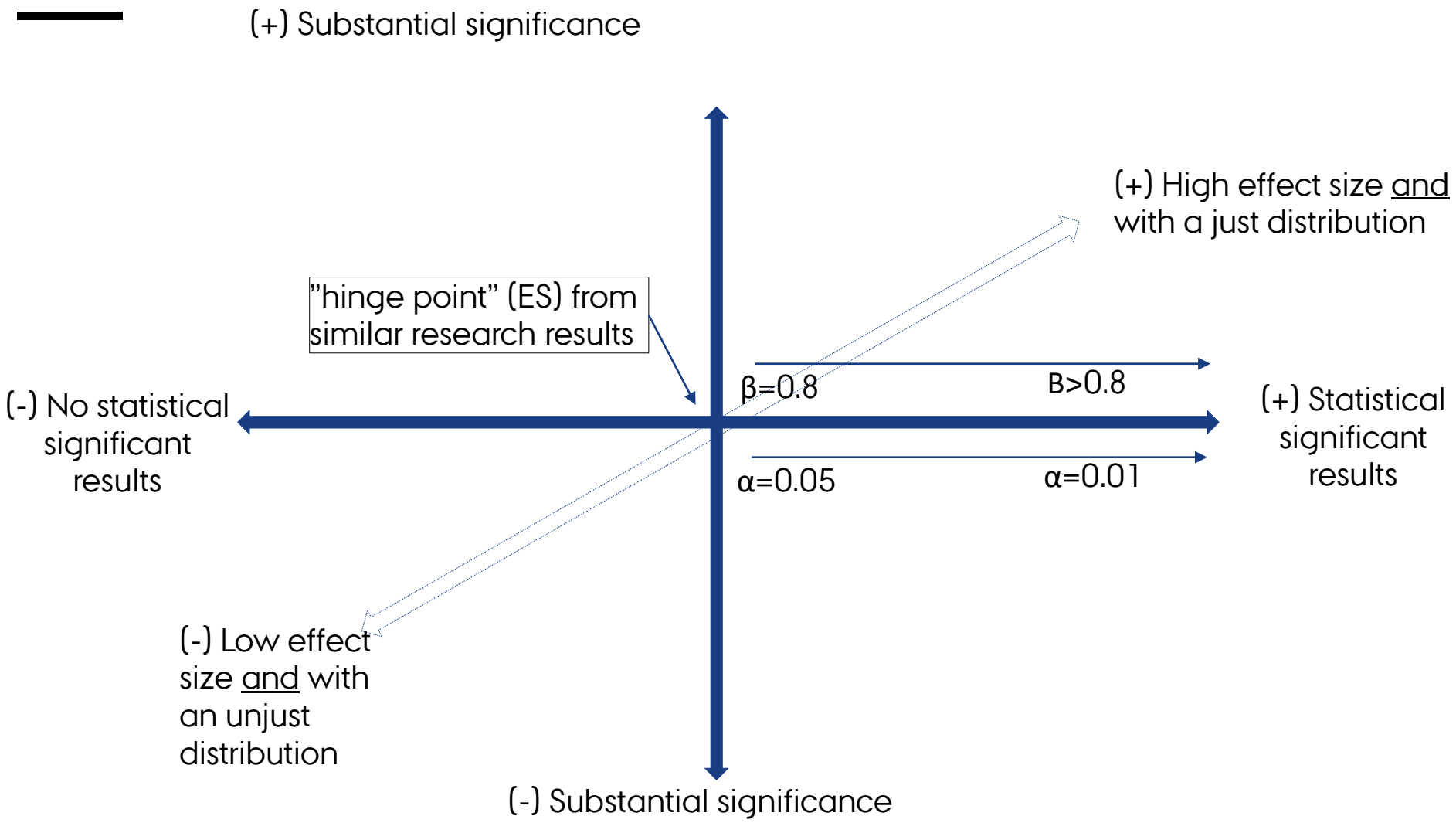
THIRD CONTINUUM WITHIN QUANTITATIVE EDUCATIONAL RESEARCH



(-) Low effect size and with an unjust distribution

(+) High effect size and with a just distribution

WHERE TO MOVE QUANTITATIVE EDUCATIONAL RESEARCH IN THE SPACE OF RESEARCH RESULT TYPES



DOES IT MAKE A SUBSTANTIAL DIFFERENCE THAT WE FOUND A STATISTICAL SIGNIFICANT DIFFERENCE?

The Danish TIMSS 2015 Report:

Kønsforskelle i præstationer	Der er små, men statistisk signifikante kønsforskelle i matematik.
	Matematik: drenge = 542 piger = 536
	Der er ingen statistisk signifikante forskelle i natur/teknologi.
	Natur /teknologi: drenge = 529 piger = 525

AARHUS
UNIVERSITY