

***Evidence for the effectiveness of
educational interventions:
Methodological challenges in meta-
analyses***

Monica Melby-Lervåg, UiO

Charles Hulme, UCL, UiO

Arne Lervåg, UiO

Take-home message

Methodology matters: If we continue to base reviews of interventions on studies with poor methodology this will lead us astray



Study:

A syntethesis of meta-analyses



Melby-Lervåg, Lervåg & Hulme, work in progress.

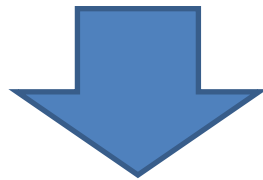
Method

Systematic search for reviews of educational interventions that have used a quantitative summary of results after 1998

The meta-analysis had to examine an intervention that could in some way inform about amelioration of difficulties related to:

Decoding, reading comprehension, language skills, mathematic skills, general learning disorders, attention/hyperactivity, other behavioral/emotional problems or bullying.

The meta-analysis had to provide a mean effect size of an academic achievement or behavioral outcome that was based on a group design (i.e. meta-analyses purely based on single case studies were excluded)

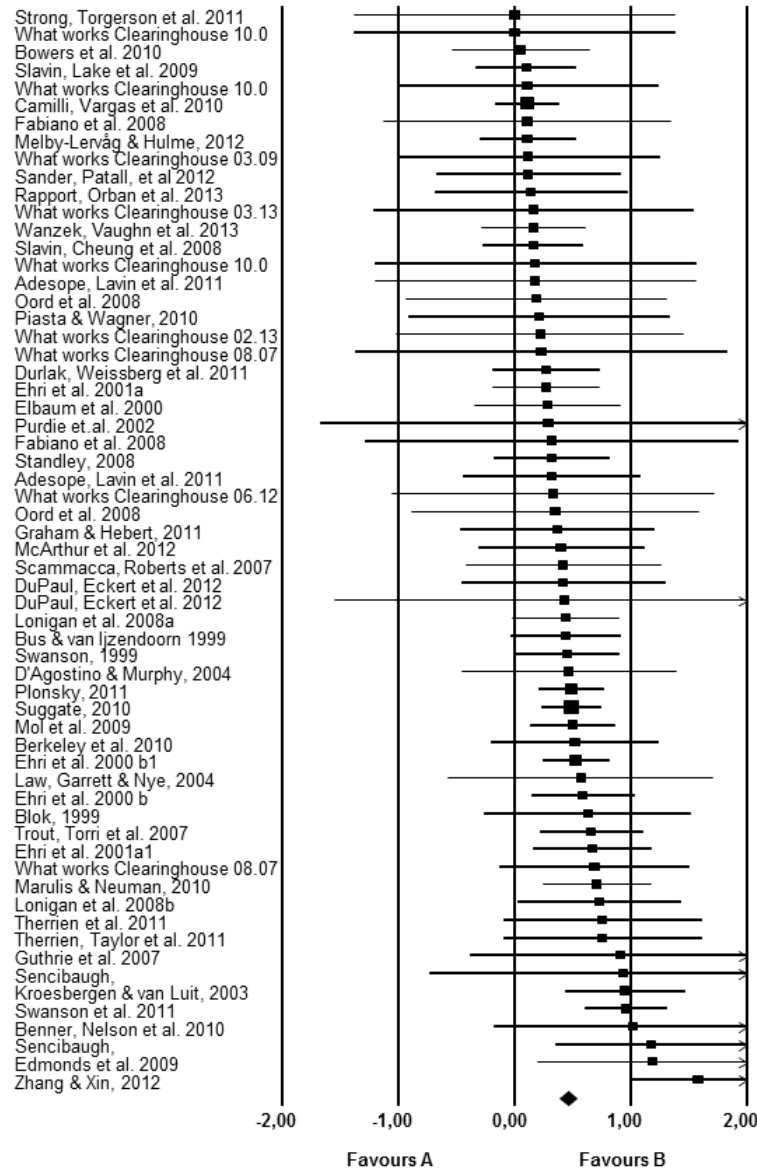


70 meta-analyses included, 3145 single studies

Results

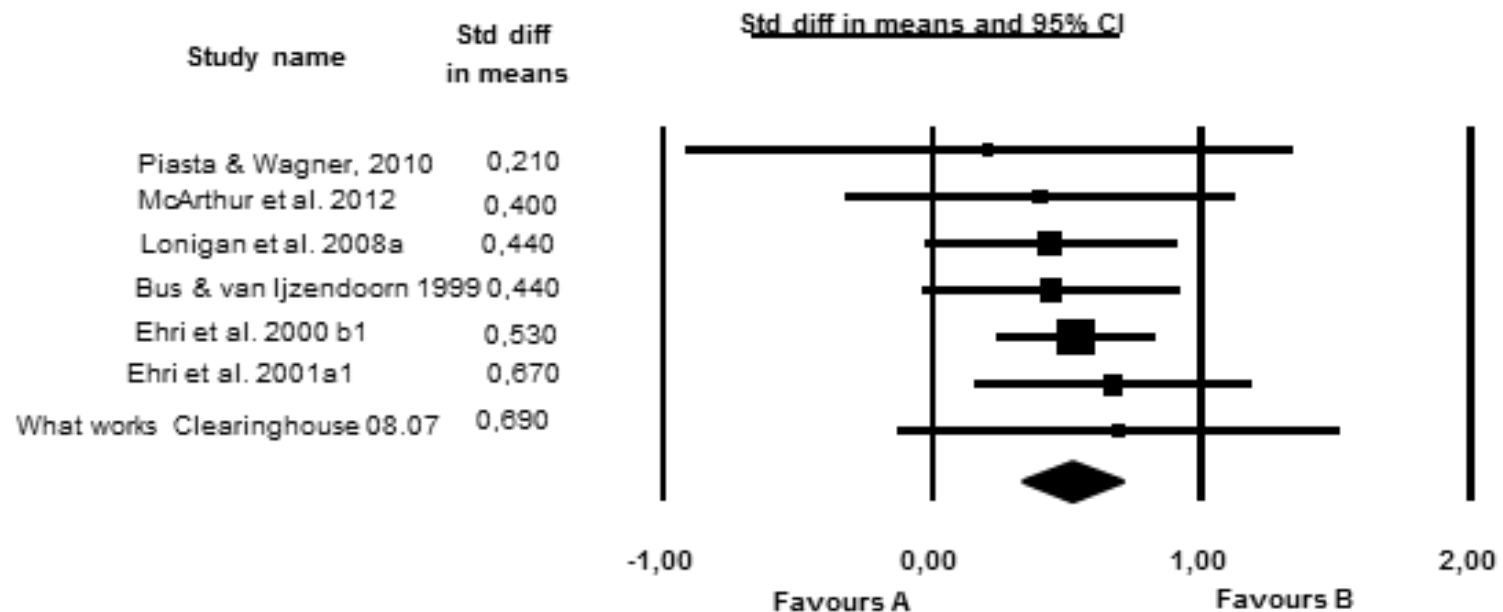


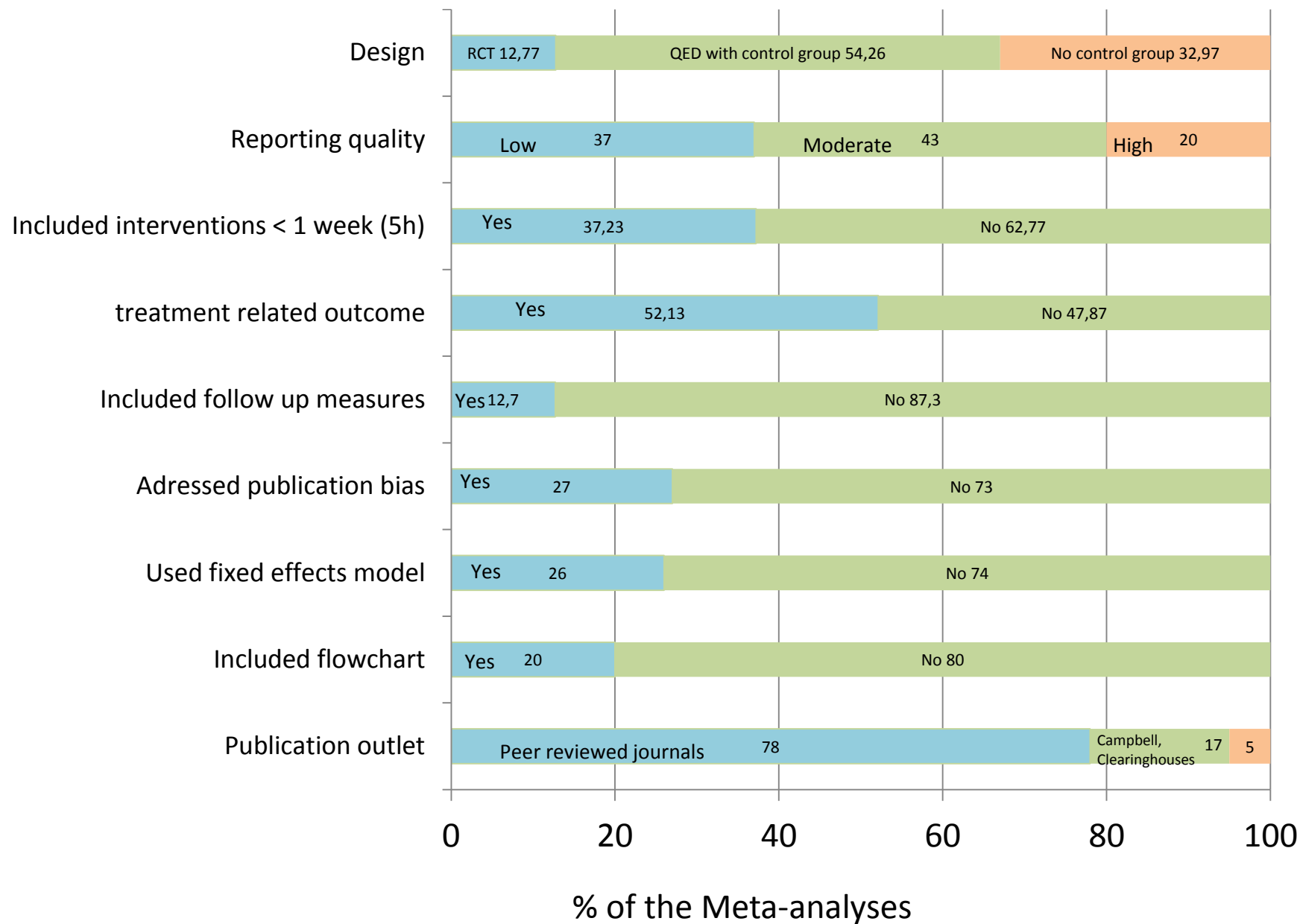
Academic achievement



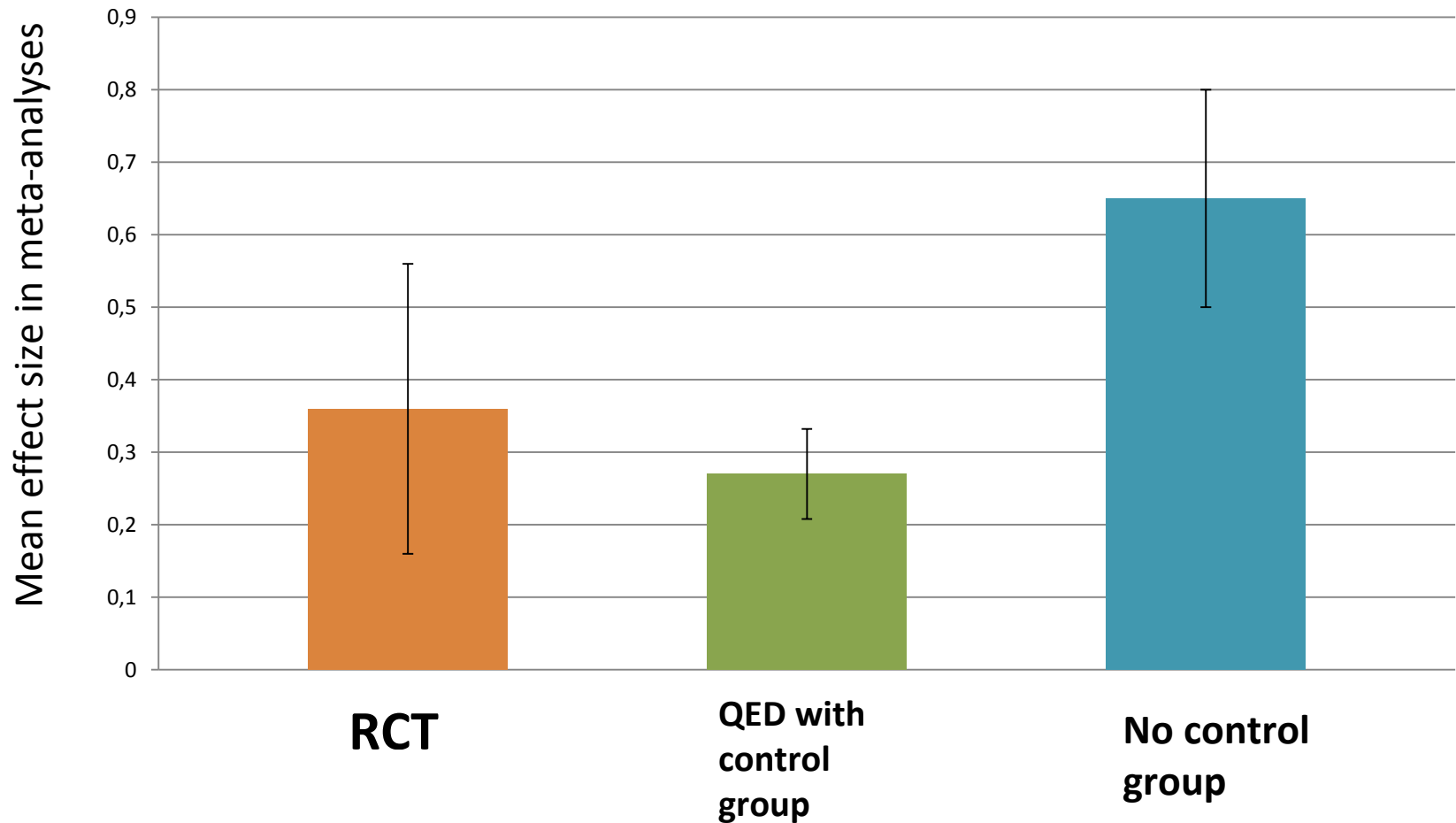
Mean effect sizes reported in the meta-analyses range from 0-1.58 standard deviation units, with a mean $d = 0.43$ (0.35-0.51)

The effect from phonological awareness training on word reading skills

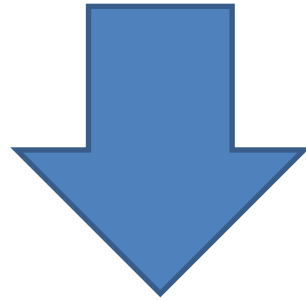




Differences in mean effect size for different designs



Only 233 of the 3145 intervention studies
were randomised controlled trials.



Serious methodological weaknesses,
studies not suited to conclude about
intervention effects

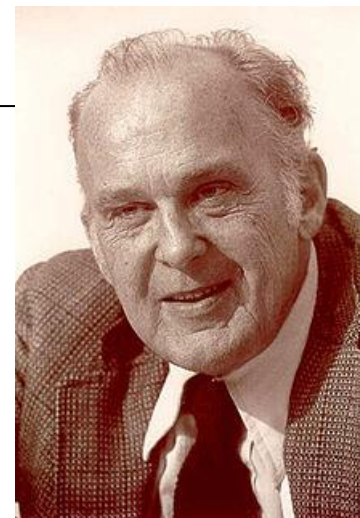
In the 20% of the studies with high reporting quality, all except one were from “What works clearinghouse”, EPPI centre, Campbell or Cochrane.

The majority of the reviews that had included studies with the highest quality (i.e. RCTS or QEDs with proper baseline control) was also conducted by these institutions

REFORMS AS EXPERIMENTS ¹

DONALD T. CAMPBELL ²

Northwestern University

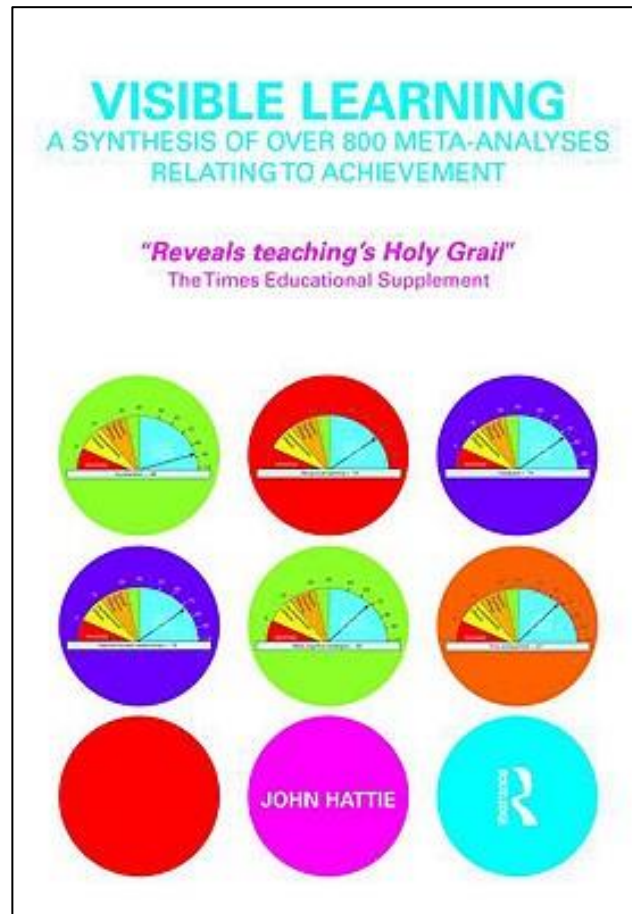


THE United States and other modern nations should be ready for an experimental approach to social reform, an approach in which we try out new programs designed to cure specific social problems, in which we learn whether or not these programs are effective, and in which we retain, imitate, modify, or discard them on the basis of apparent effectiveness on the multiple imperfect criteria available. Our readiness for this stage is indicated by the inclusion of specific provisions for program evaluation in the first wave of the "Great Society" legislation, and by the current congressional proposals for establishing "social indicators" and socially relevant "data banks." So long have we had good intentions in this regard that many may feel we are already at this stage, that we already are continuing or discontinuing programs on the basis of assessed effectiveness. It is a theme of this article that this is not at all so, that most ameliorative programs end up with *no* interpretable evaluation (Etzioni, 1968; Hyman & Wright, 1967; Schwartz, 1961). We must look hard at the sources of this condition, and design

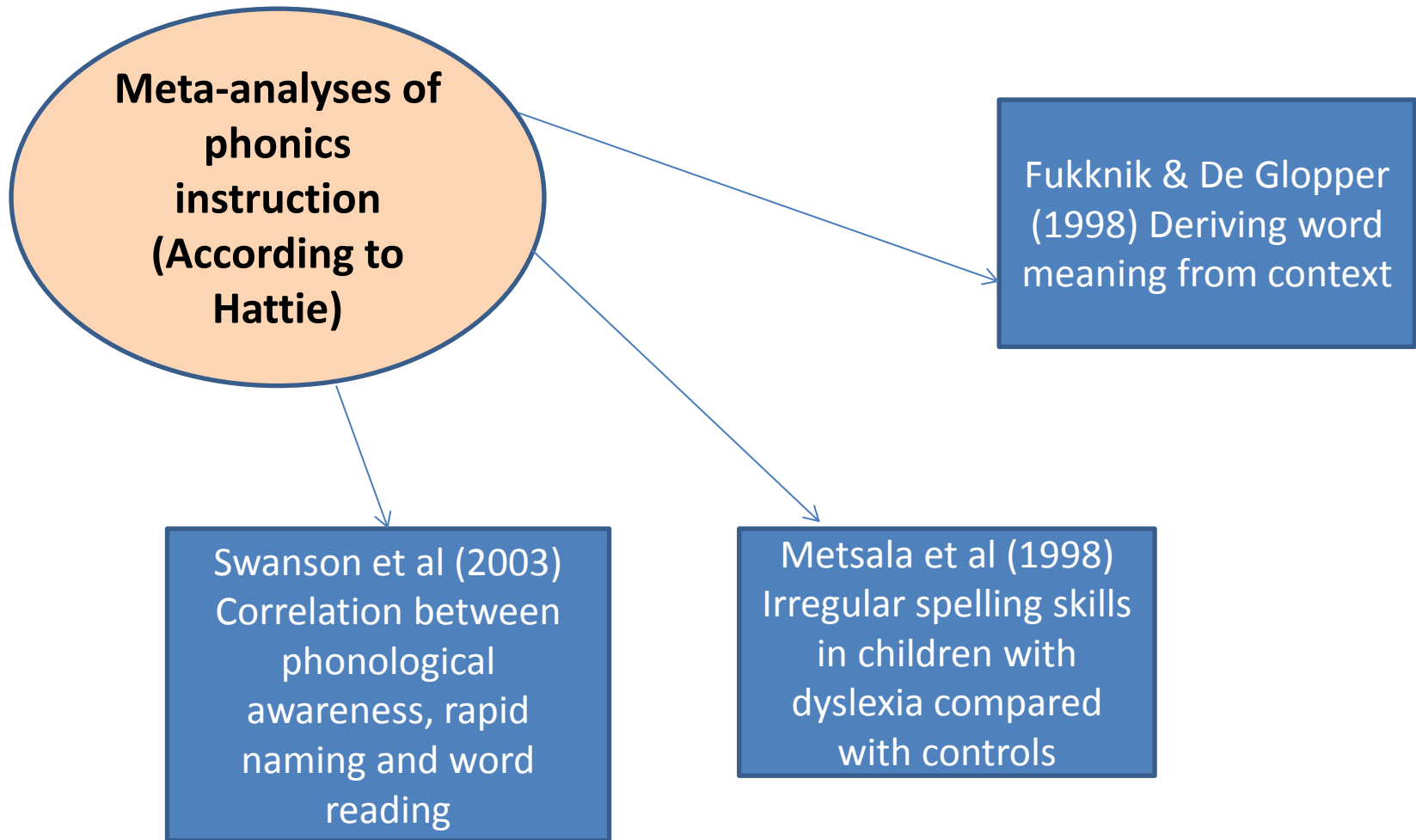
decisions are made in a political arena, and involve political jeopardies that are often sufficient to explain the lack of hard-headed evaluation of effects. Removing reform administrators from the political spotlight seems both highly unlikely, and undesirable even if it were possible. What is instead essential is that the social scientist research advisor understand the political realities of the situation, and that he aid by helping create a public demand for hard-headed evaluation, by contributing to those political inventions that reduce the liability of honest evaluation, and by educating future administrators to the problems and possibilities.

For this reason, there is also an attempt in this article to consider the political setting of program evaluation, and to offer suggestions as to political postures that might further a truly experimental approach to social reform. Although such considerations will be distributed as a minor theme throughout this article, it seems convenient to begin with some general points of this political nature.

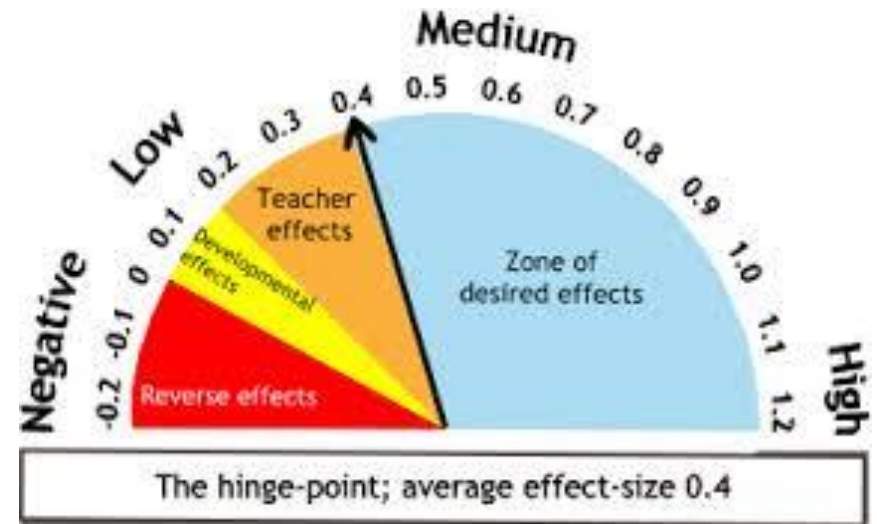
A more famous attempt of synthesising meta-analyses.....



Example based on the Hattie book



Hattie's effect barometer



Setting 0.4 as a minimum effect size means that the results from the majority of the RCTs will be dismissed.



From Visible Learning (2009)



Take-home message

Methodology matters: If we continue to base reviews of interventions on studies with poor methodology this will lead us astray



Thank you for the attention!

Foto: Kathrine Nordli, «Airborne»