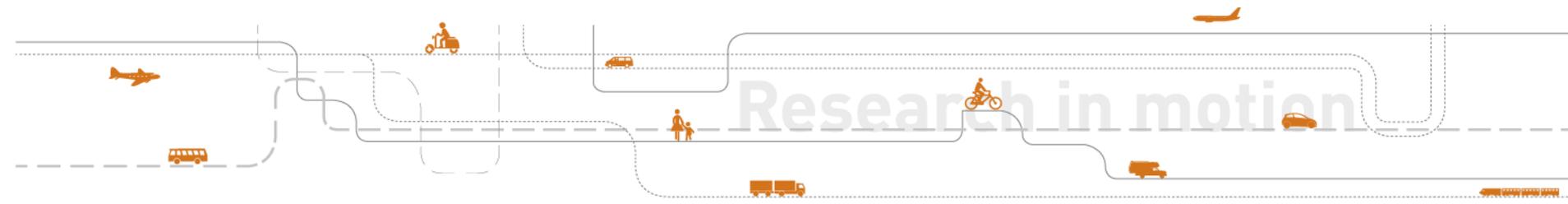


Methods and common pitfalls to gaining valid knowledge

ICoRSI International Dialogues on Road Safety

Rune Elvik, Institute of Transport Economics
(re@toi.no)



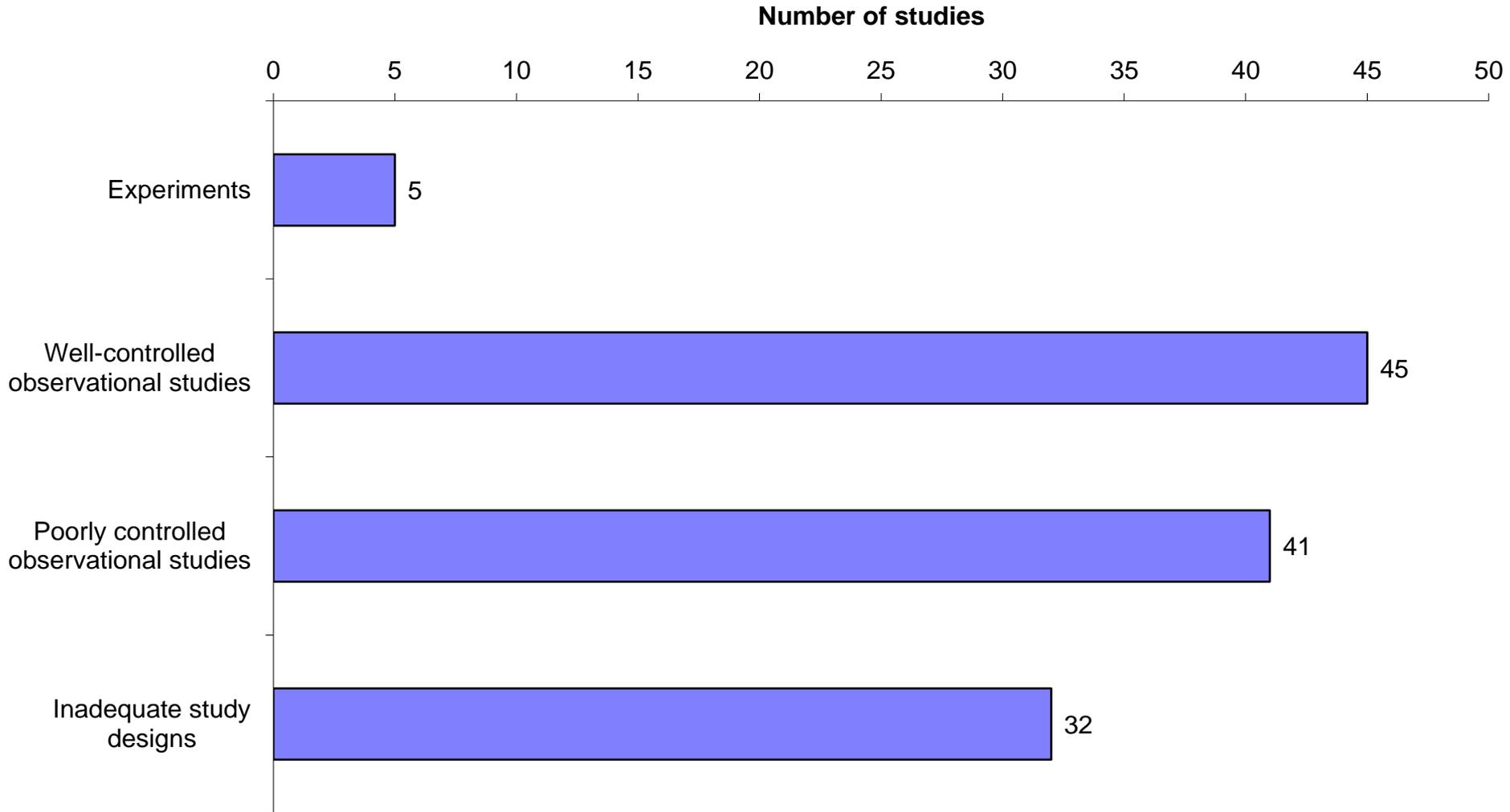
What is valid knowledge?

- Knowledge is valid if we cannot give reasons for doubting its truth
- We cannot claim to know the truth
- We can only claim that we do not (for the moment) have reasons to believe that a study did not find the truth
- Knowledge is valid if it approximates the truth
- To judge validity, we need to specify what we mean by validity and what influences it

The validity framework of Shadish, Cook and Campbell

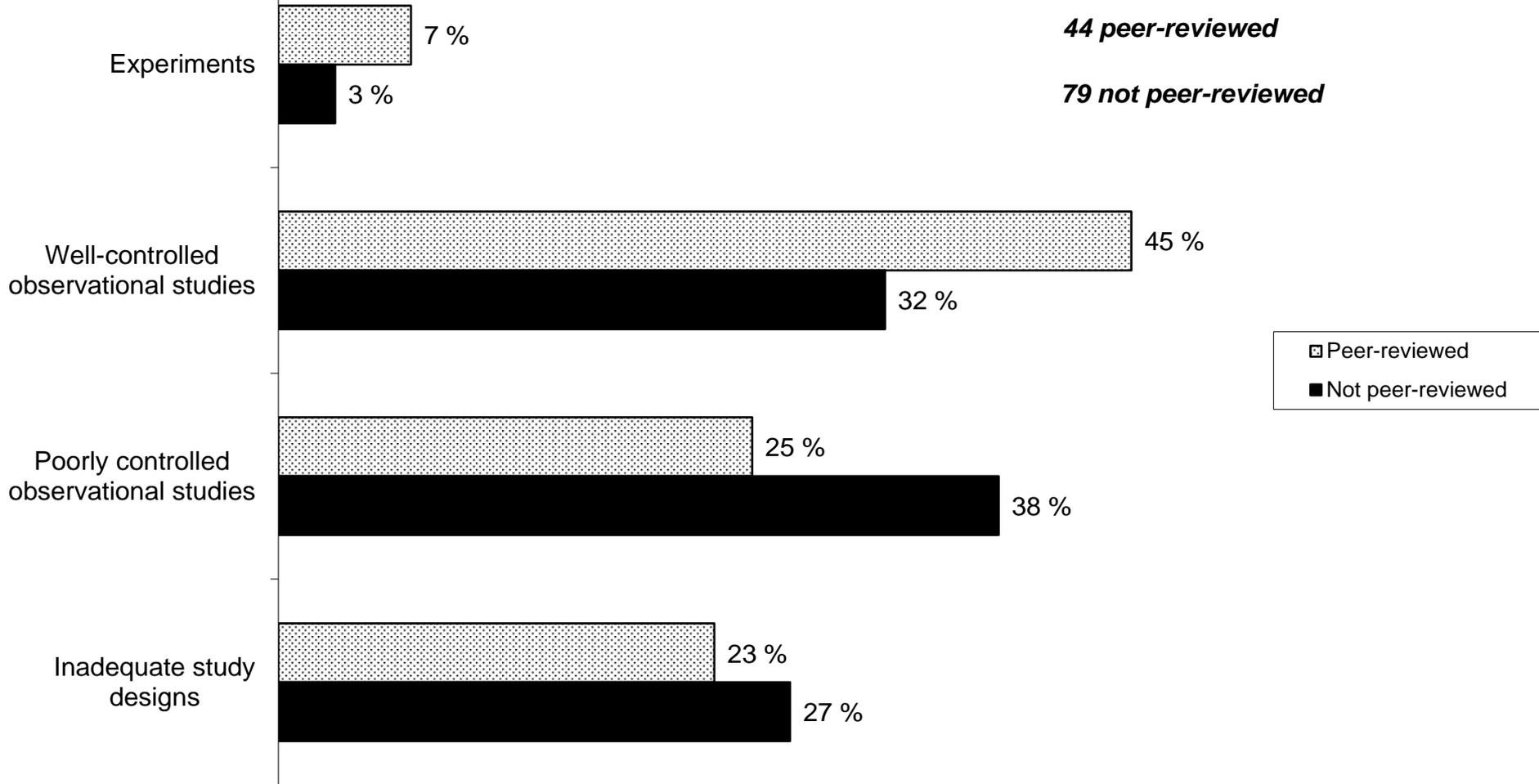
- Four types of validity:
 - *Statistical conclusion validity*
 - *Construct (theoretical) validity*
 - *Internal validity*
 - *External validity*
- This typology can be applied to road safety studies
- The most valid knowledge is gained from randomised, controlled trials (experiments) performed in large samples
- Any other method is considerably less valid

123 road safety evaluation studies according to design



Percentage of evaluation studies

0 % 10 % 20 % 30 % 40 % 50 %



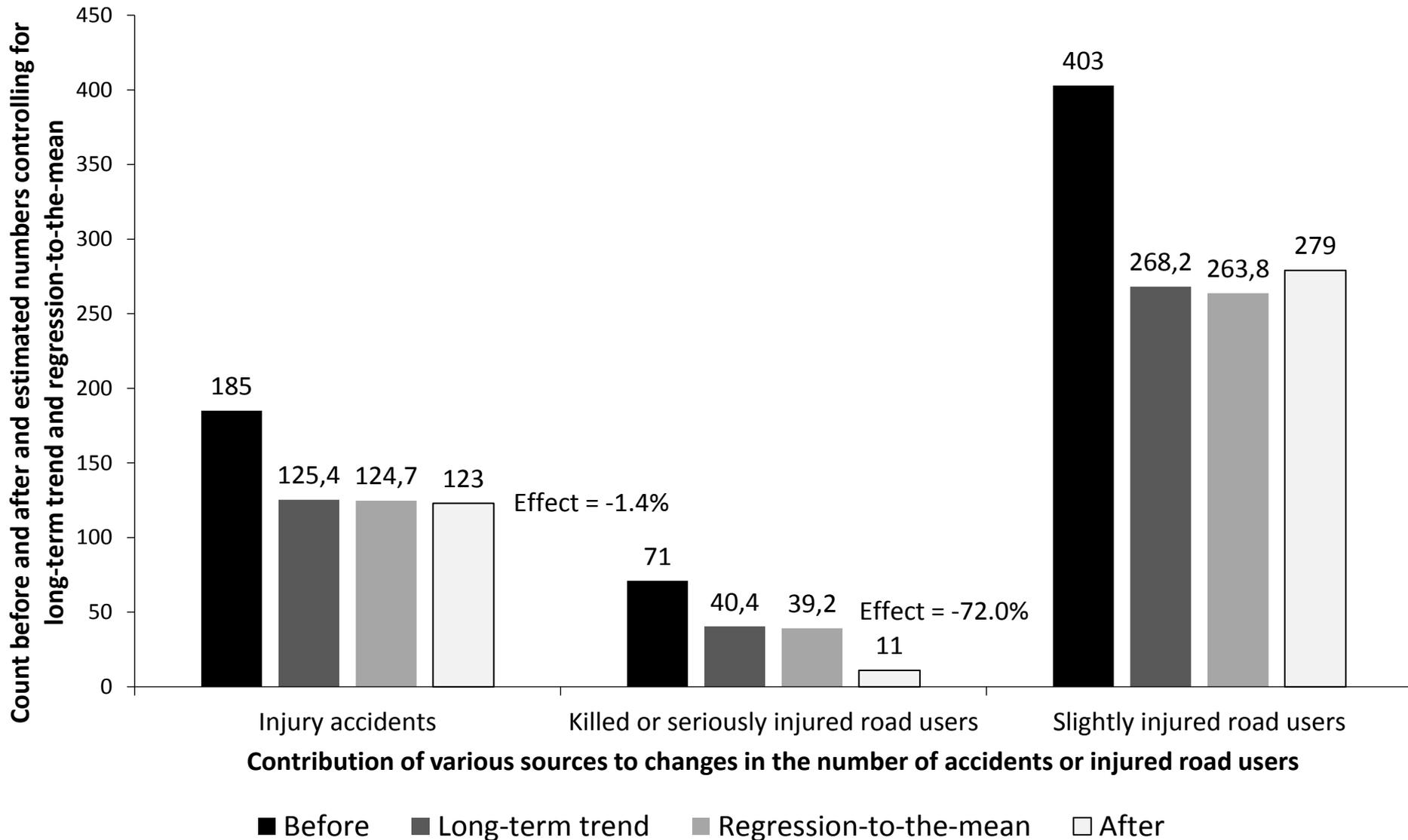
Our dilemma in a nutshell

- What we tell policymakers:
- ***We strongly recommend that you base your policy on the best available scientific evidence***
- What we ought to add:
- ***But we must warn you that most of this evidence should not be trusted***

Is there a way out of the dilemma?

- It will never be the case that all, or even most, road safety evaluation studies are experimental
- Observational studies will continue to dominate
- We should trust observational before-after studies that control for at least well-known confounding factors
- We should trust well-controlled case-control studies
- We should even trust a set of poorly controlled studies if their findings as a whole «make sense»
- A single poorly controlled study should not be trusted

Factors contributing to changes in road safety - new motorway



Making sense of poor studies – the case of road lighting

- The most common study design:

Condition	Before	After	Total
Darkness	40	25	65
Daylight	80	75	155
Total	120	100	220

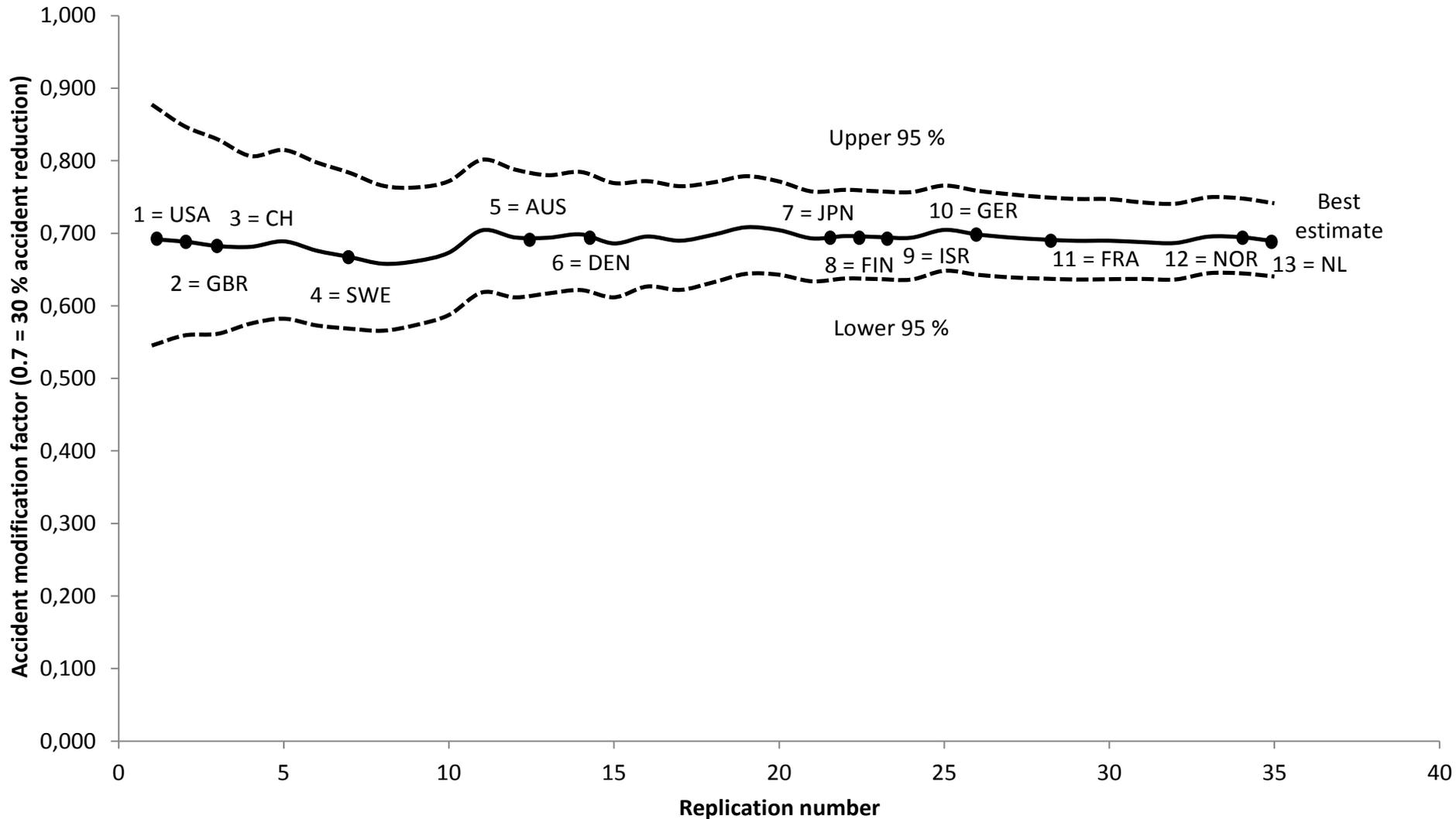
- Odds ratio estimate of effect = $(25/40)/(75/80) = 0.667$
- There is no control for:
 - *Long-term trends*
 - *Regression-to-the-mean*
 - *Changes in traffic volume*

The whole set of studies of road lighting makes sense

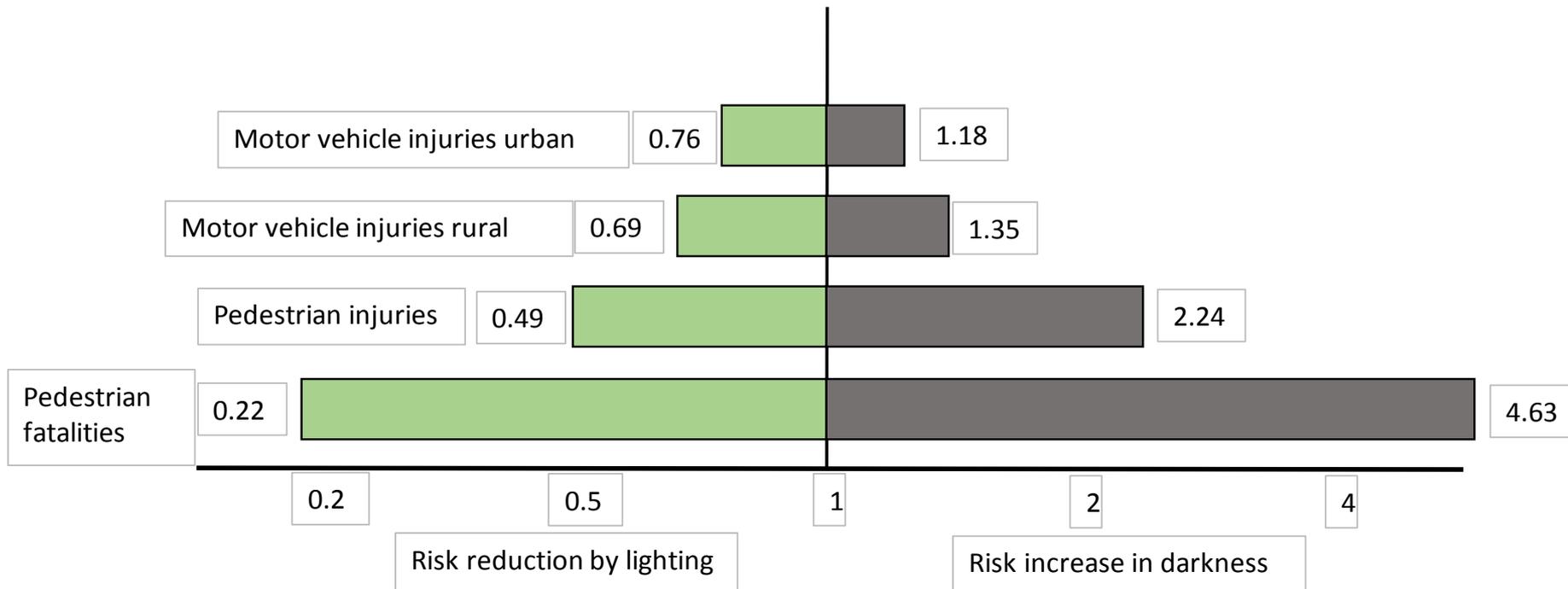
- Darkness is a universal risk factor; one would therefore expect the results of evaluation studies to be consistent over time and between countries
- One would expect a smaller effect in urban areas, where there are many sources of artificial lighting, than in the countryside
- One would expect a larger effect on fatal crashes than on less serious crashes

- Meta-analysis can be used to test these hypotheses

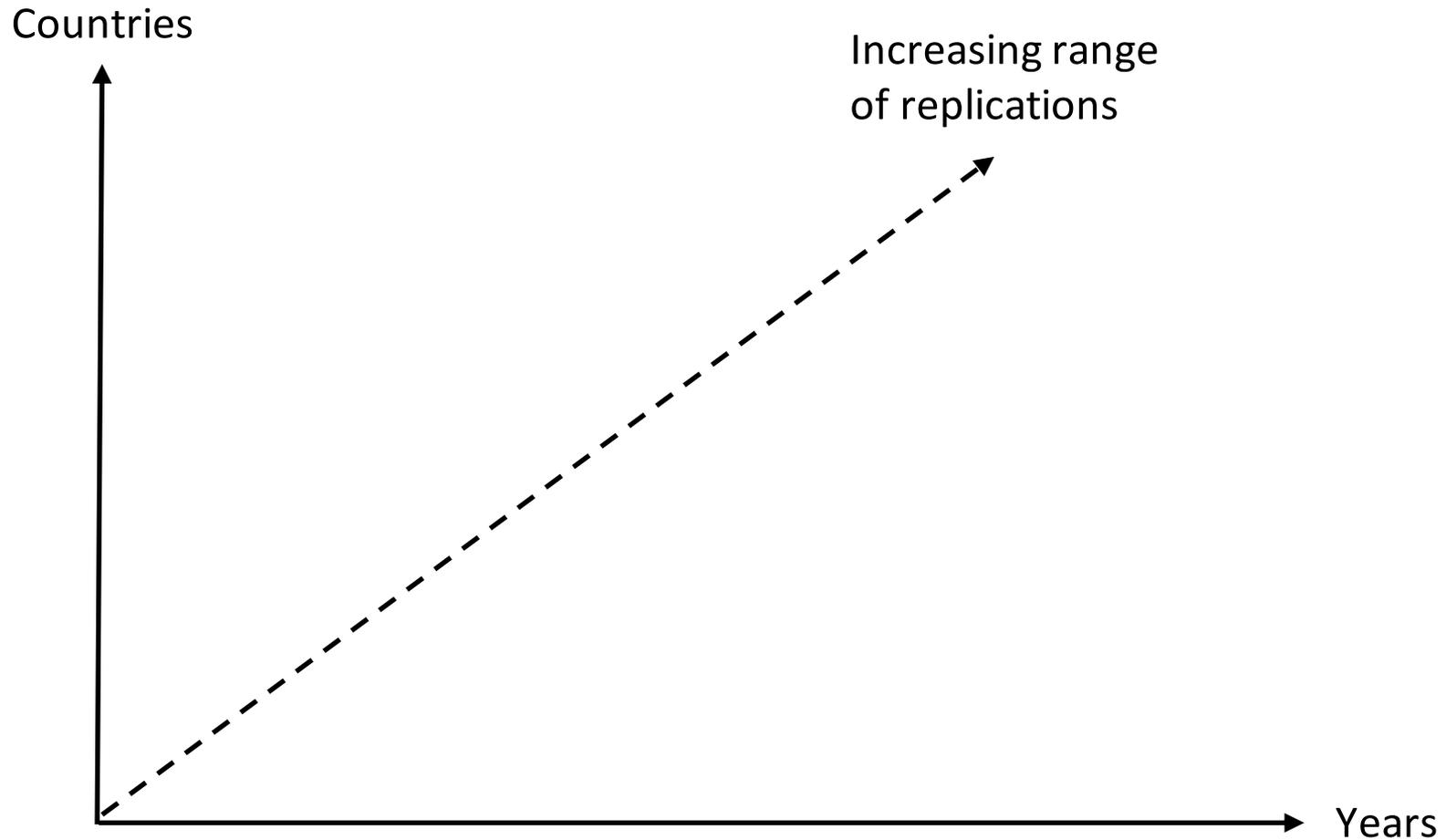
The effects of road lighting are the same in all countries where they have been evaluated



Relationship of effects to risks



The range of replications



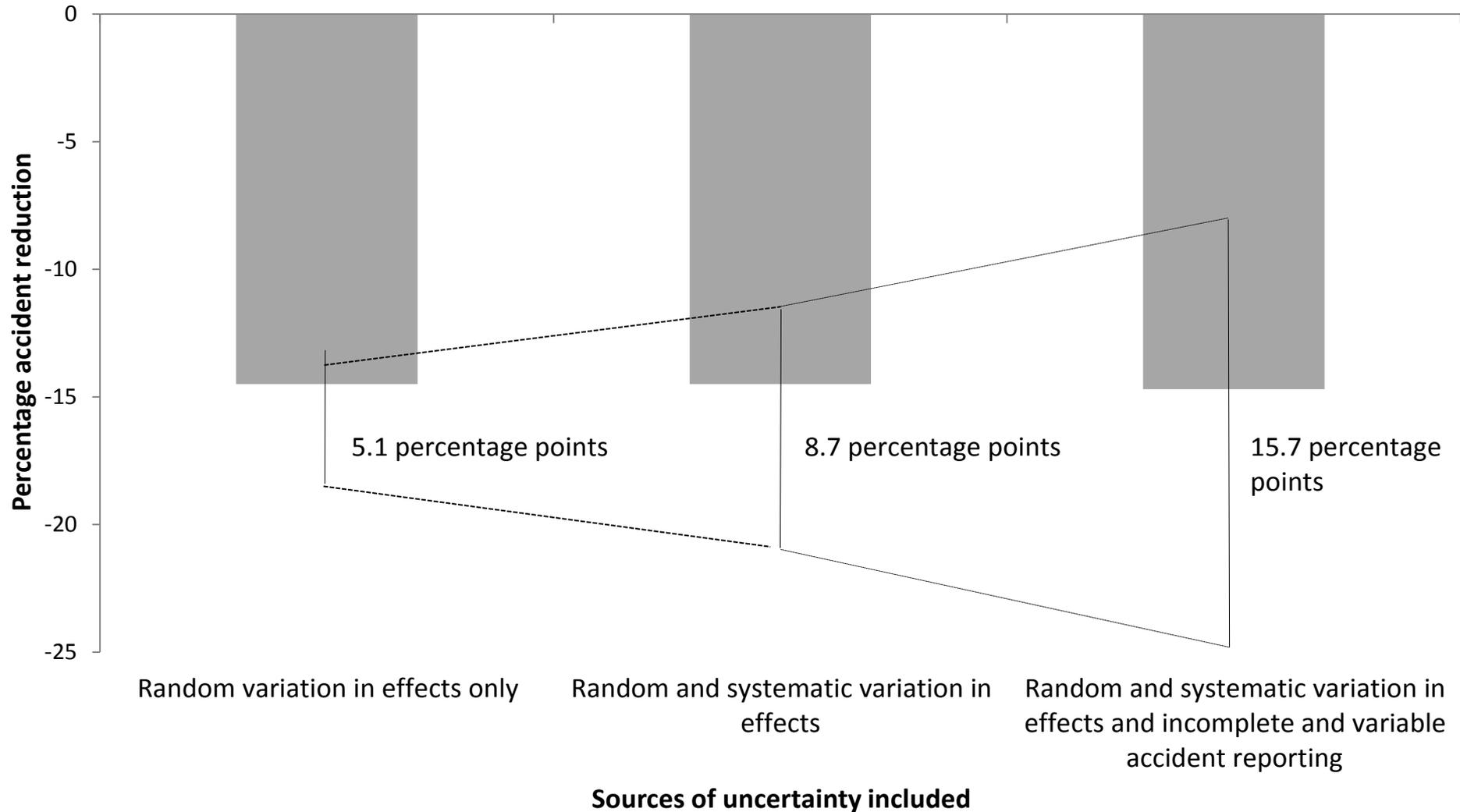
Consistent and theoretically plausible evidence

- Is available for a number of important risk factors and road safety measures:
 - *Road lighting*
 - *Roundabouts*
 - *Guardrails*
 - *Seat belts*
 - *Motor cycle helmets*
 - *Speed and speed limits*
- In all these cases the effects are plausible by reference to the laws of physics
- How about road safety measures primarily targeting the human factor?

Measures targeting human factors

- There are randomised controlled trials consistently showing that driver training is ineffective
- I believe this is a theoretically plausible finding
- The literature on campaigns is a mess – lots of poor studies
- Attempts to make sense of these studies have, I think, been unsuccessful
- Enforcement? It is probably effective in many cases

Contribution of three sources of uncertainty to summary estimates of effect of traffic calming on injury accidents



Can pitfalls be avoided?

- We can do before-after studies that control for long-term trends and regression-to-the-mean
- Case-control studies are notoriously prone to confounding, but should be trusted if their findings are theoretically plausible
- Incomplete reporting of crashes or injuries is a potentially important source of bias and imprecision which it is difficult to control for
- But that problem should have a technical solution that should be implemented