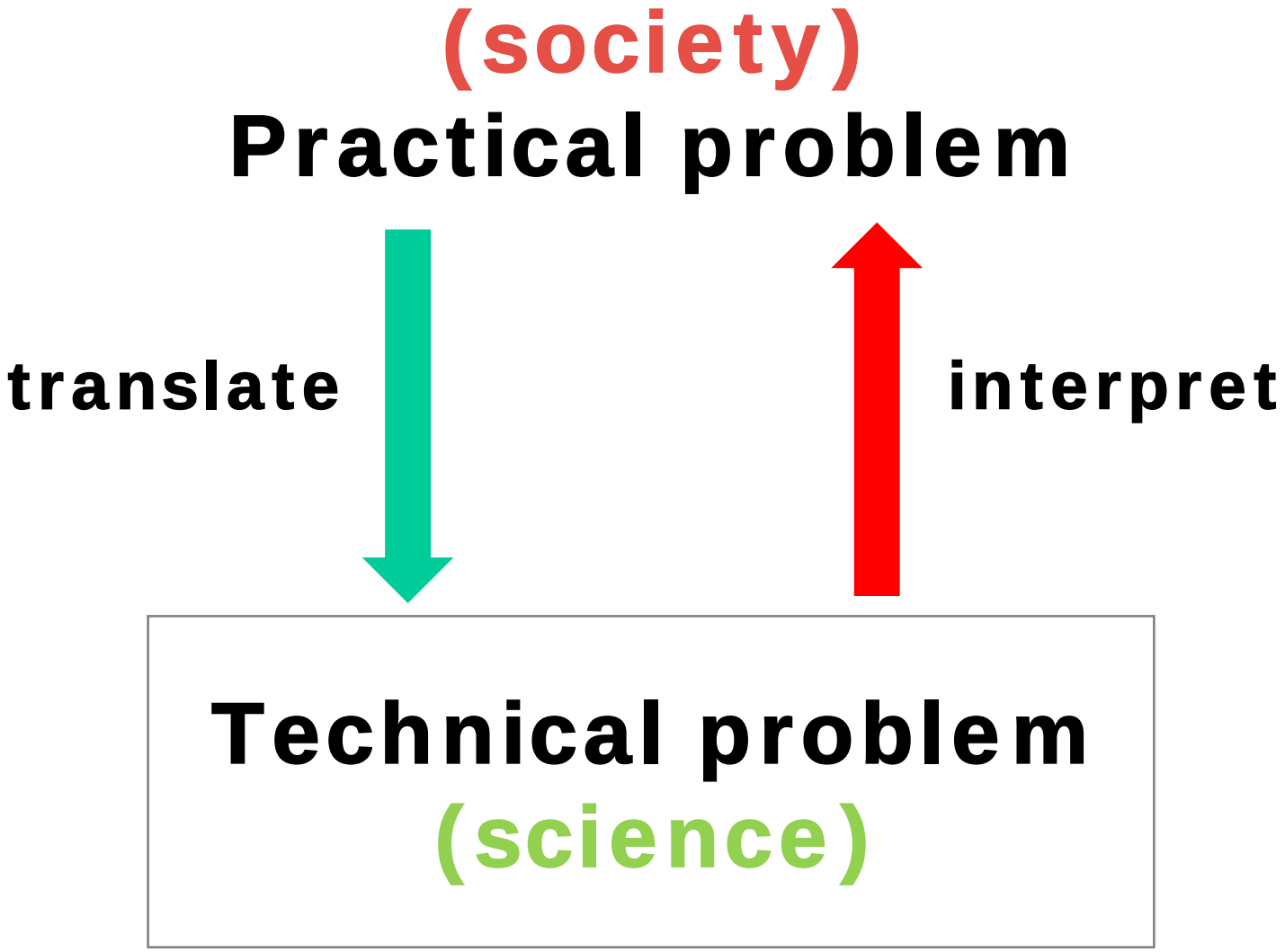
Knowledge Quality Assessment
an intro to The Guidance & NUSAP

Jeroen P. van der Sluijs



Ravetz, J., 1971, Scientific Knowledge and its Social Problems, Oxford University Press.

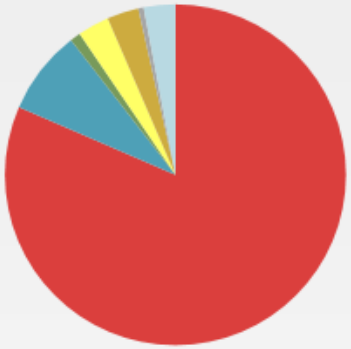
The pitfall of *lamp-posting* in translating *practical problems* into *technical problems*

Searching where the light shines
may not help to solve the practical problem…

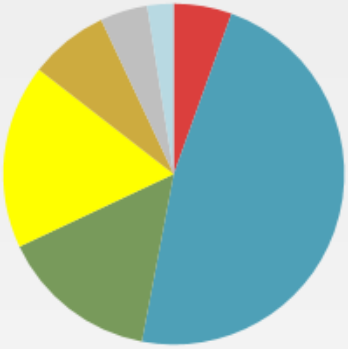

LISTEN TO
THE SCIENCE
BAN

Whose science counts?

Beekeepers

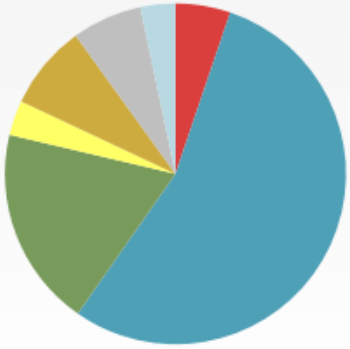


Authorisation board

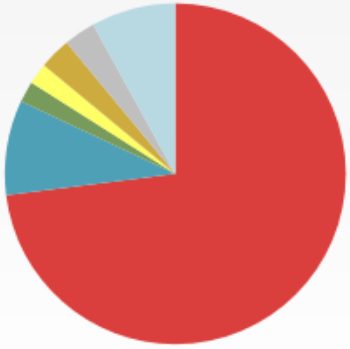


Ministry of Agriculture

Bayer



Academic researchers

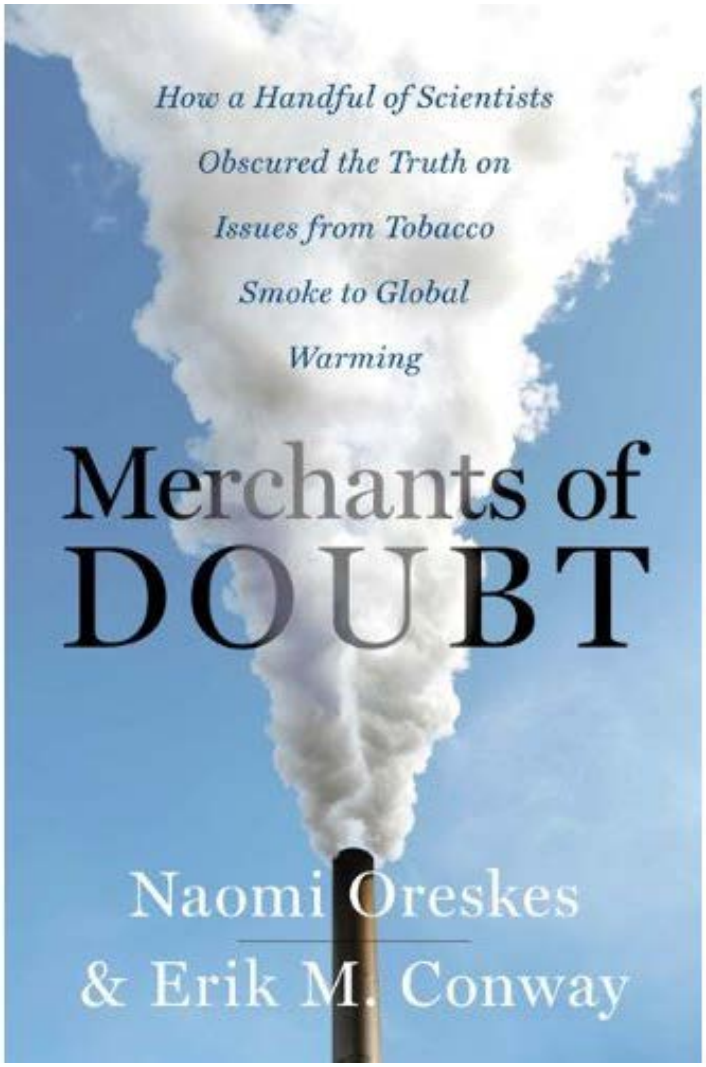


- Neonic pesticdes
- Diseases
- Pollen
- Sunflower strains
- Climate
- Genetics

(Maxim & Van der Sluijs, 2010)

3 sources that fuel dissent in scientific community

Conflicts of interests



Institutionalized practices

Published online 5 October 2011 | *Nature* **478**, 7 (2011) | doi:10.1038/478007a

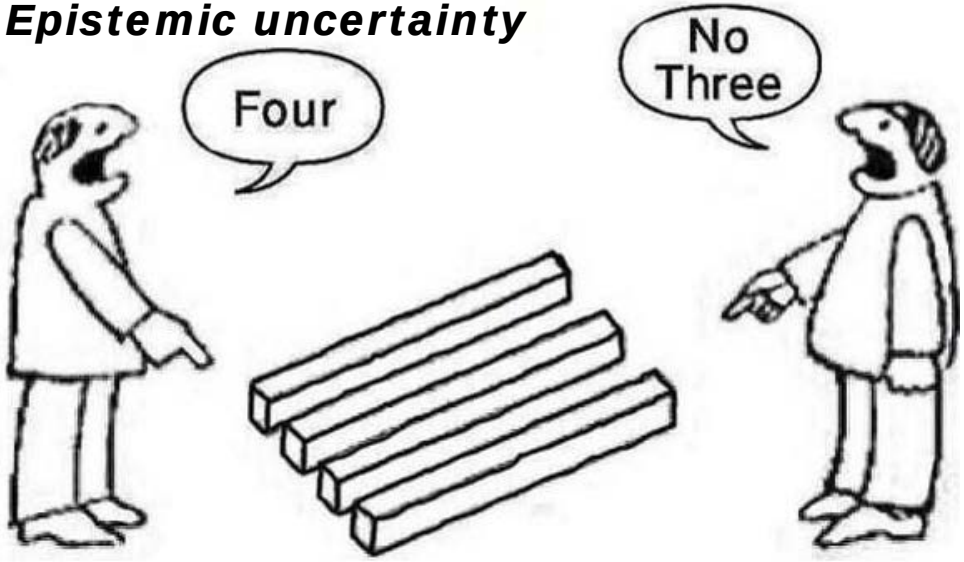
Column: World View

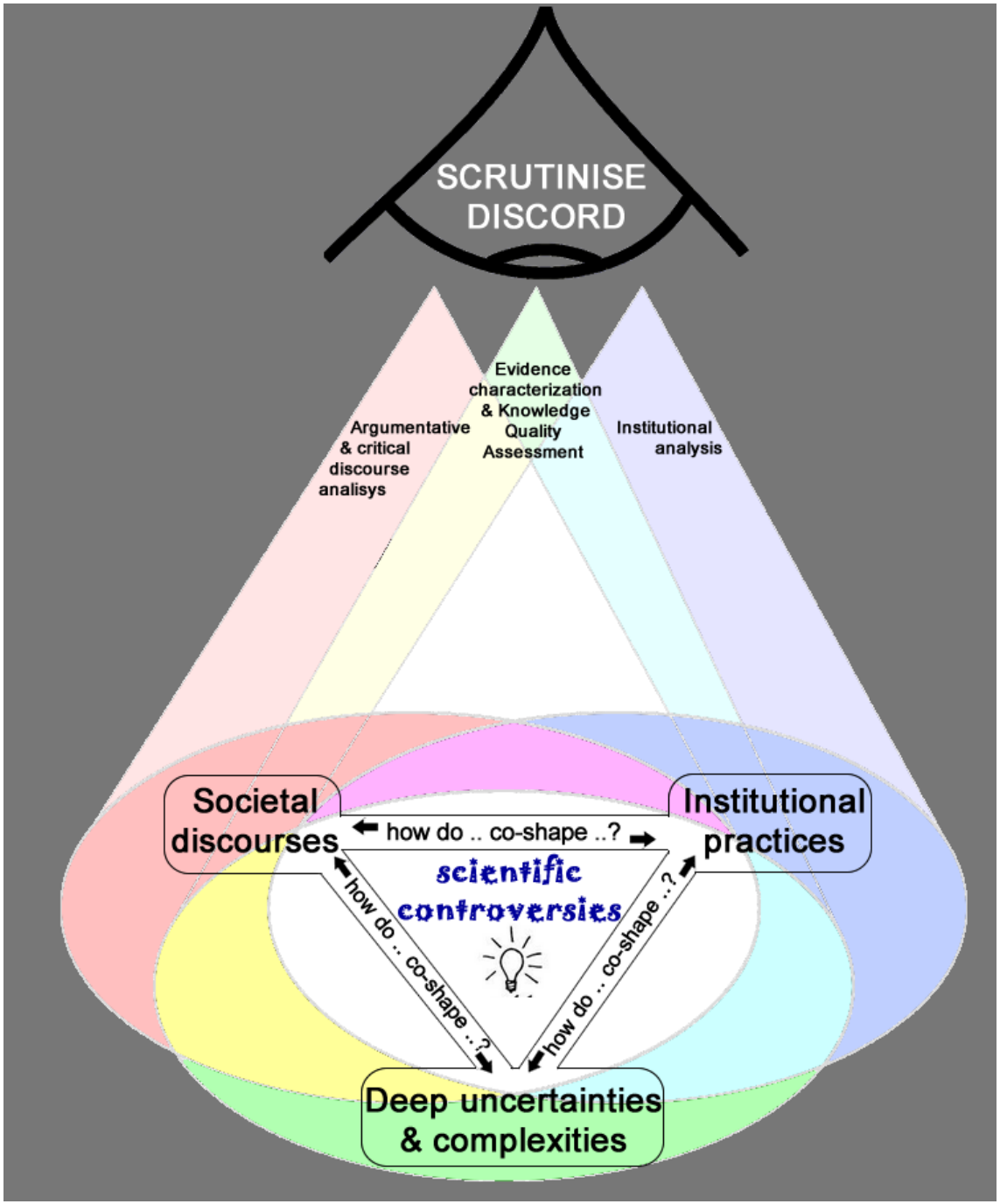
The voice of science: let's agree to disagree

Consensus reports are the bedrock of science-based policy-making. But disagreement and arguments are more useful, says Daniel Sarewitz.

Daniel Sarewitz

Epistemic uncertainty





New way of looking at scientific controversies

"By shining light on its dynamics from 3 different perspectives (**discourse analysis, evidence characterization**, **institutional analysis**) it seeks to reveal how 3 key factors (**deep uncertainties; societal discourses; institutional practices**) co-shape one another to produce the typical patterns that can be observed in scientific controversies."

Van der Sluijs, 2014

Clark & Majone 1985

Critical Appraisal of Scientific Inquiries with Policy Implications

1. Criticism by whom?

Critical roles

- Scientist
- Peer group
- Program Manager or Sponsor
- Policy maker
- Public interests groups

Clark & Majone 1985

Criticism of what?

Critical modes:

- Input
  - data; methods, people, competence, (im)matureness of field
- Output
  - problem solved? hypothesis tested?
- Process
  - good scientific practice, procedures for review, documenting etc.

Table 1. Critical criteria.

(Clark & Majone, 1985)

Critical Role	Critical Mode: Input	Critical Mode: Output	Critical Mode: Process
Scientist	Resource and time constraints; available theory; institutional support; assumptions; quality of available data; state of the art.	Validation; sensitivity analyses; technical sophistication; degree of acceptance of conclusions; impact on policy debate; imitation; professional recognition.	Choice of methodology (e.g., estimation procedures); communication; implementation; promotion; degree of formalization of analytic activities within the organization.
Peer Group	Quality of data; model and/or theory used; adequacy of tools; problem formulation. Input variables well chosen? Measure of success specified in advance?	Purpose of the study. Are conclusions supported by evidence? Does model offend common sense? Robustness of conclusions; adequate coverage of issues.	Standards of scientific and professional practice; documentation; review of validation techniques; style; interdisciplinarity.
Program Manager or Sponsor	Cost; institutional support within user organization; quality of analytic team; type of financing (e.g., grant vs. contract).	Rate of use; type of use (general education, program evaluation, decisionmaking, etc.); contribution to methodology and state of the art; prestige. Can results be generalized, applied elsewhere?	Dissemination; collaboration with users. Has study been reviewed?
Policymaker	Quality of analysts; cost of study; technical tools used (hardware and software). Does problem formulation make sense?	Is output familiar and intelligible? Did study generate new ideas? Are policy indications conclusive? Are they consonant with accepted ethical standards?	Ease of use; documentation. Are analysts helping with implementation? Did they interact with agency personnel? With interest groups?
Public Interest Groups	Competence and intellectual integrity of analysts. Are value systems compatible? Problem formulation acceptable? Normative implications of technical choices (e.g., choices of data).	Nature of conclusions; equity. Is analysis used as rationalization or to postpone decision? All viewpoints taken into consideration? Value issues.	Participation; communication of data and other information; adherence to strict rules of procedure.

Clark & Majone 1985

Meta quality criteria:

- Adequacy
  - reliability, reproducibility, uncertainty analysis etc.
- Value
  - Internal: how well is the study carried out?
  - External: fitness for purpose, fitness for function
  - Personal: subjectivity, preferences, choices, assumptions, bias
- Effectiveness
  - Does it help to solve practical problems
- Legitimacy
  - numinous: natural authority, independence, credibility, competence
  - civil: agreed procedures

Scandal at the Netherlands Environmental Assessment Agency

RIVM / De Kwaadsteniet (1999)

"RIVM over-exact prognoses based on virtual reality of computer models"

Newspaper headlines:

- Environmental institute lies and deceits
- Fuss in parliament after criticism on environmental numbers
- The bankruptcy of the environmental numbers
- Society has a right on fair information, RIVM does not provide it

NL Environmental Assessment Agency (RIVM/MNP) Guidance: Systematic reflection on uncertainty & quality in:

Foci	Key issues
Problem framing	Other problem views; interwovenness with other problems; system boundaries; role of results in policy process; relation to previous assessments
Involvement of stakeholders	Identifying stakeholders; their views and roles; controversies; mode of involvement
Selection of indicators	Adequate backing for selection; alternative indicators; support for selection in science, society, and politics
Appraisal of knowledge base	Quality required; bottlenecks in available knowledge and methods; impact of bottlenecks on quality of results
Mapping and assessing relevant uncertainties	Identification and prioritisation of key uncertainties; choice of methods to assess these; assessing robustness of conclusions
Reporting uncertainty information	Context of reporting; robustness and clarity of main messages; policy implications of uncertainty; balanced and consistent representation in progressive disclosure of uncertainty information; traceability and adequate backing

Problem framing and context

- Explore rival problem frames
- Relevant aspects / system boundary
- Typify problem structure
- Problem lifecycle / maturity
- Role of study in policy process
- Uncertainty in socio-political context

Type-III error:
Assessing the wrong problem by incorrectly accepting the false meta-hypothesis that there is no difference between the boundaries of a problem, as defined by the analyst, and the actual boundaries of the problem (Dunn, 1997).

Context validation (Dunn, 1999).
The validity of inferences that we have estimated the proximal range of rival hypotheses.

Context validation can be performed by a participatory bottom-up process to elicit from scientists and stakeholders rival hypotheses on causal relations underlying a problem and rival problem definitions.

Involvement of stakeholders

- Identify relevant stakeholders.
- Identification of areas of agreement and disagreement among stakeholders on value dimensions of the problem.
- Recommendations on when to involve different stakeholders in the assessment process.

Roles of stakeholders

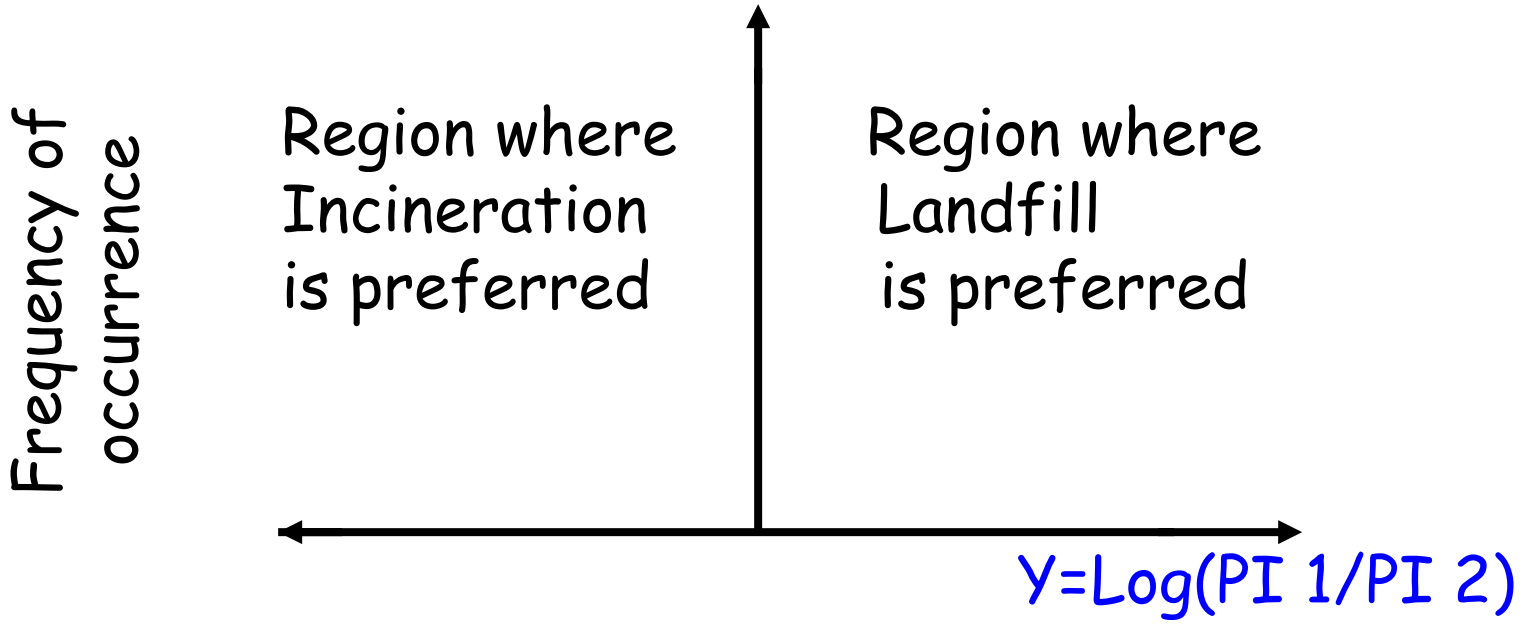
- (Co-) definer of the problems to be addressed
  - What knowledge is relevant?
- Source of knowledge
- Quality control of the science (for instance: review of assumptions)

Indicators

- How well do indicators used address key aspects of the problem?
- Use of proxies
- Alternative indicators?
- Limitations of indicators used?
- Scale and aggregation issues
- Controversies in science and society about these indicators?

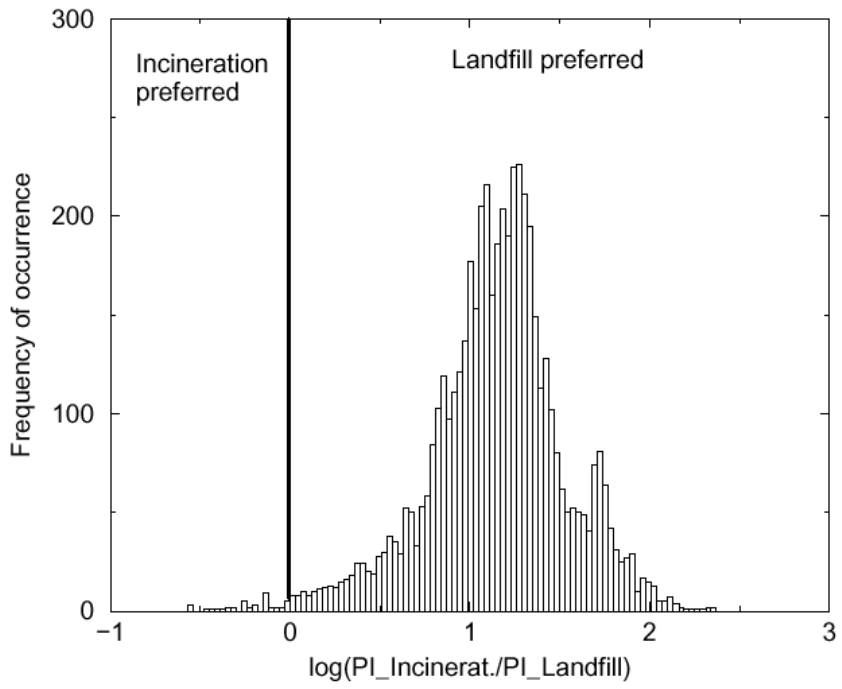
High uncertainty is not the same as low quality

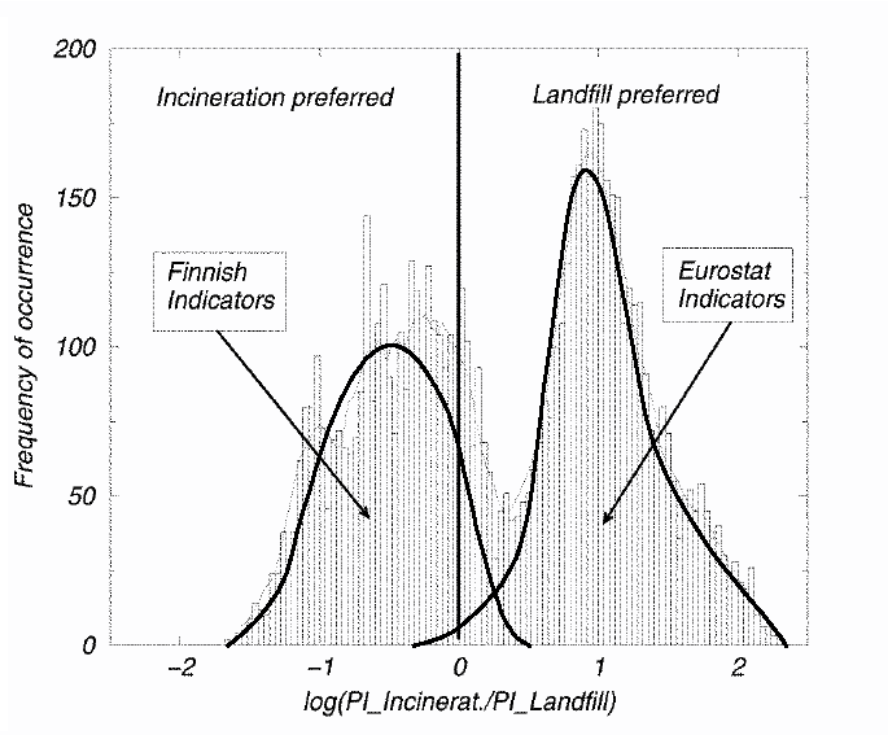
Example: imagine the inference is Y = the logarithm of the ratio between the two pressure-on-decision indices PI1 and PI2



High uncertainty is not the same as low quality,

but..... methodological uncertainty can de dominant





(slide borrowed from Andrea Saltelli)

Do we know enough to quantify?

Risbey & Kandlikar (2007): What format is in accordance with the level of knowledge on the quantity?

- Full probability density function
  - Robust, well defended distribution
- Bounds
  - Well defended percentile bounds
- First order estimates
  - Order of magnitude assessment
- Expected sign or trend
  - Well defended trend expectation
- Ambiguous sign or trend
  - Equally plausible contrary trend expectations
- Effective ignorance
  - Lacking or weakly plausible expectations

Uncertainty is more than a number

Dimensions of uncertainty:

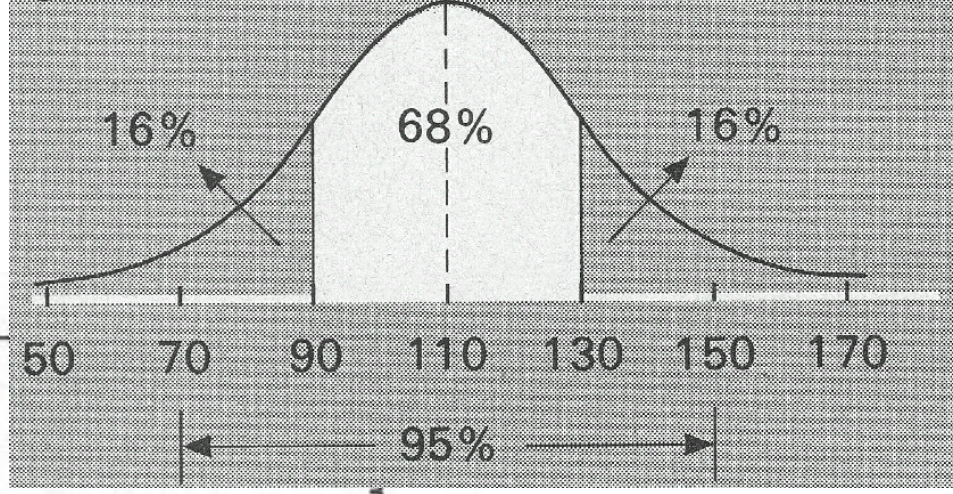
- Technical (inexactness)
- Methodological (unreliability)
- Epistemological (ignorance)
- Societal (limited social robustness)

Reliability intervals normal distributions

$\pm \sigma = 68\ \%$

$\pm 2\sigma = 95\ \%$

$\pm 3\sigma = 99.7\ \%$



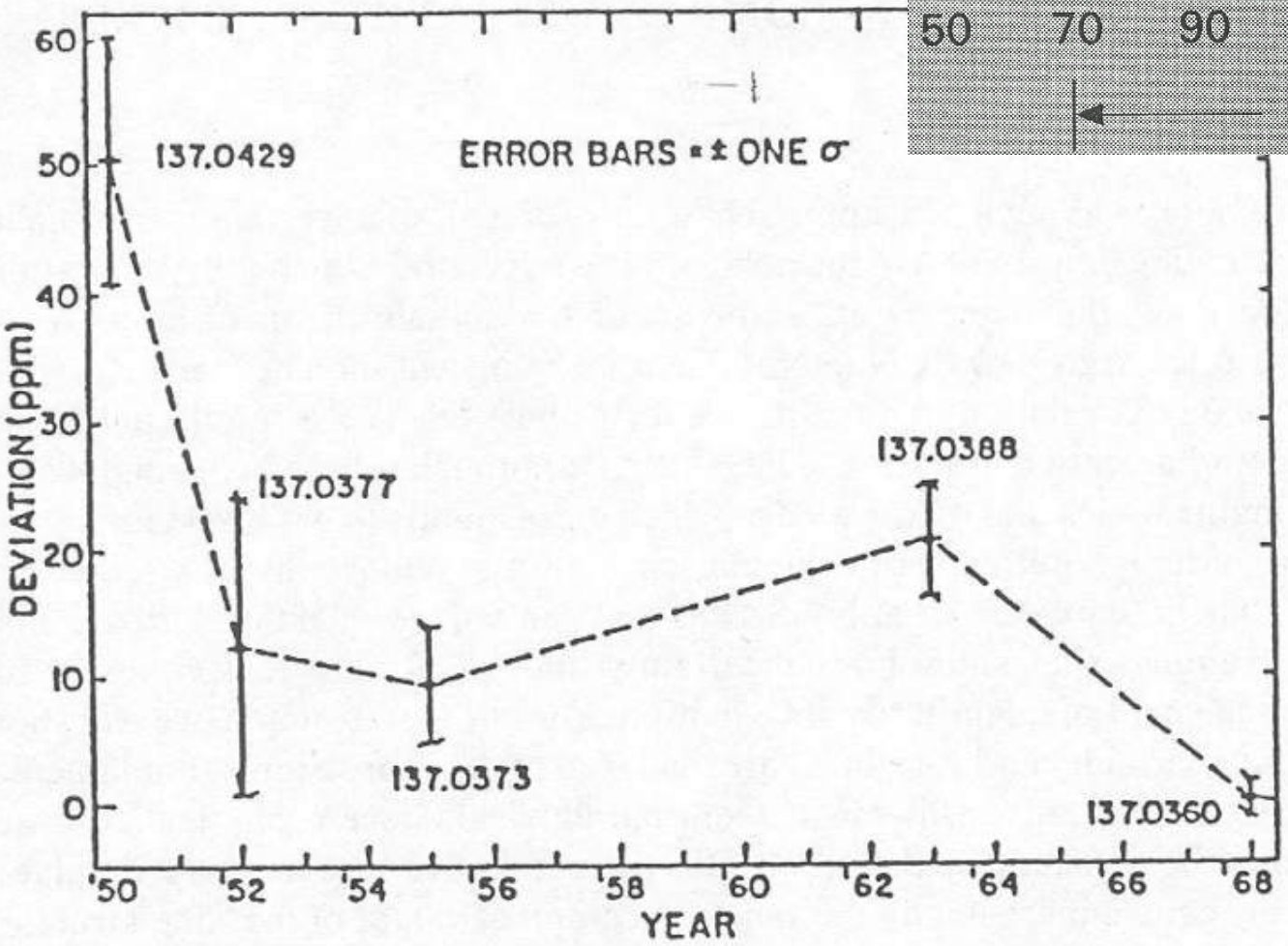
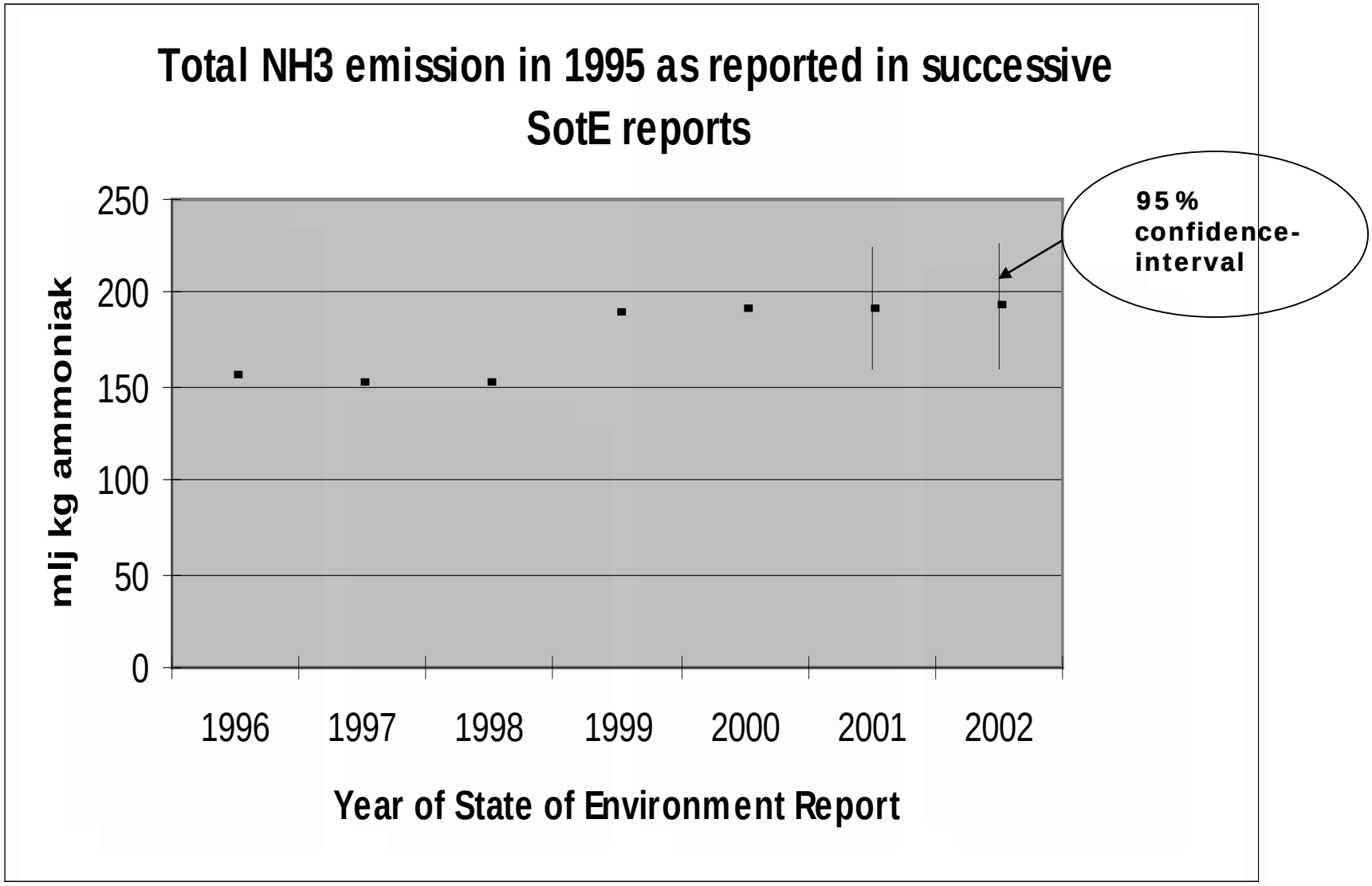


Fig. 1. Successive recommended values of the fine-structure constand α^{-1} (B. N. Taylor *et al.*, 1969, 7)


Total NH3 emission in 1995 as reported in successive SotE reports
mlj kg ammoniak
250
200
150
100
50
0
1996
1997
1998
1999
2000
2001
2002
Year of State of Environment Report
95% confidence-interval

NUSAP: Qualified Quantities

Classic scientific notational system:

- **N**umeral **U**nit **S**pread

For problems in the post-normal domain, add two qualifiers:

- **A**ssessment & **P**edigree

  "Assessment" expresses expert judgement on reliability of numeral + spread

  "Pedigree" expresses multi-criteria evaluation of the strength of a number by looking at:

  - Background history by which the number was produced
  - Underpinning and scientific status of the number

Example Pedigree matrix parameter strength

Code	Proxy	Empirical	Theoretical basis	Method	Validation
4	Exact measure	Large sample direct mmts	Well established theory	Best available practice	Compared with indep. mmts of same variable
3	Good fit or measure	Small sample direct mmts	Accepted theory partial in nature	Reliable method commonly accepted	Compared with indep. mmts of closely related variable
2	Well correlated	Modeled/derived data	Partial theory limited consensus on reliability	Acceptable method limited consensus on reliability	Compared with mmts not independent
1	Weak correlation	Educated guesses / rule of thumb est	Preliminary theory	Preliminary methods unknown reliability	Weak / indirect validation
0	Not clearly related	Crude speculation	Crude speculation	No discernible rigour	No validation

Example Pedigree results

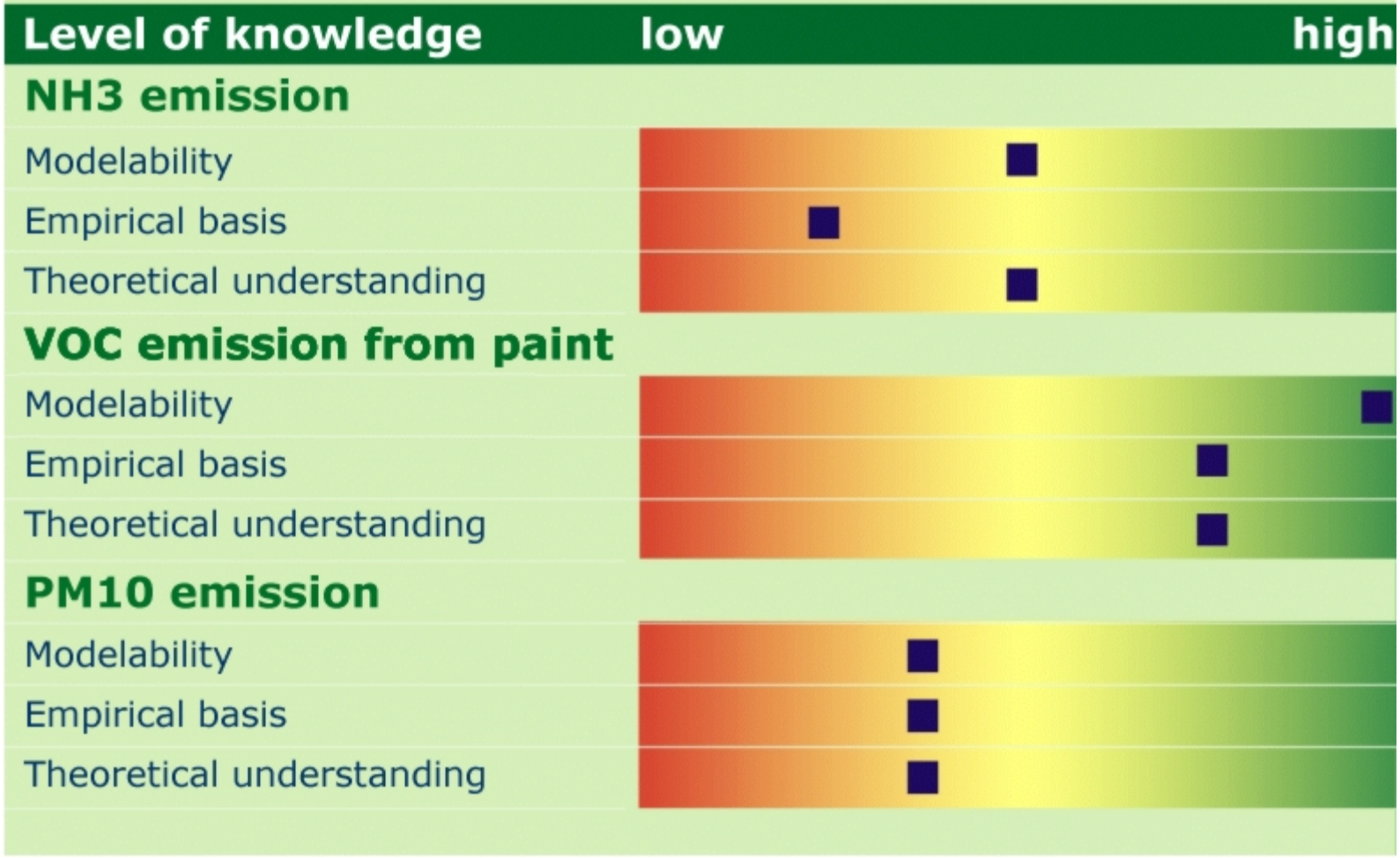
	Proxy	Empirical	Method	Validation	Strength
NS-SHI	3	3.5	4	0	0.66
NS-B&S	3	3.5	4	0	0.66
NS-DIY	2.5	3.5	4	3	0.81
NS-CAR	3	3.5	4	3	0.84
NS-IND	3	3.5	4	0.5	0.69
Th%-SHI	2	1	2	0	0.31
Th%-B&S	2	1	2	0	0.31
Th%-DIY	1	1	2	0	0.25
Th%-CAR	2	1	2	0	0.31
Th%-IND	2	1	2	0	0.31
VOS % import	1	2	1.5	0	0.28
Attribution import	1	1	2	0	0.25

Trafic-light analogy <1.4 red; 1.4-2.6 amber; >2.6 green

This example is the case of VOC emissions from paint in the Netherlands, calculated from national sales statistics (NS) in 5 sectors (Ship, Building & Steel, Do It Yourself, Car refinishing and Industry) and assumptions on additional thinner use (Th%) and a lump sum for imported paint and an assumption for its VOC percentage. See full research report on www.nusap.net for details.

Example: Air Quality

■ The position reflects the level of knowledge



In summary, NUSAP

- Has a strong theoretical foundation in the theory of knowledge and the philosophy of science
- Addresses all three dimensions of uncertainty: technical (inexactness), methodological (unreliability) and epistemological (border with ignorance) in an coherent way
- Provides a systematic framework for synthesising qualitative and quantitative assessments of uncertainty
- Can act as a bridge between the quantitative mathematical disciplines and traditions and the qualitative discursive and participatory disciplines and traditions in the field of uncertainty management.
- Helps to focus research efforts on the potentially most problematic model components
- Pinpoints specific weaknesses in these components
- Provides those who produce, use and are affected by policy-relevant knowledge a tool for a critical self-awareness of their engagement with that knowledge. It thereby fosters extended peer review processes.

UNCERTAINTY MATRIX			Level of uncertainty *(from determinism, through probability and possibility, to ignorance)*			Nature of uncertainty		Qualification of knowledge base (backing)			Value-ladenness of choices		
Location ↓			**Statistical uncertainty (range+ chance)**	**Scenario uncertainty (range as 'what-if' option)**	**Recognized ignorance**	**Knowledge-related uncertainty**	**Variability-related uncertainty**	**Weak** –	**Fair** 0	**Strong** +	**Small** –	**Medium** 0	**Large** +
Context	Ecological, technological, economic, social and political representation												
Expert judgement		Narratives; storylines; advices											
Model	**Model structure**	Relations											
	Technical model	Software & hardware implementation											
	Model parameters												
	Model inputs	Input data; driving forces; input scenarios											
Data (in general sense)	Measurements; monitoring data; survey data												
Outputs		Indicators; statements											

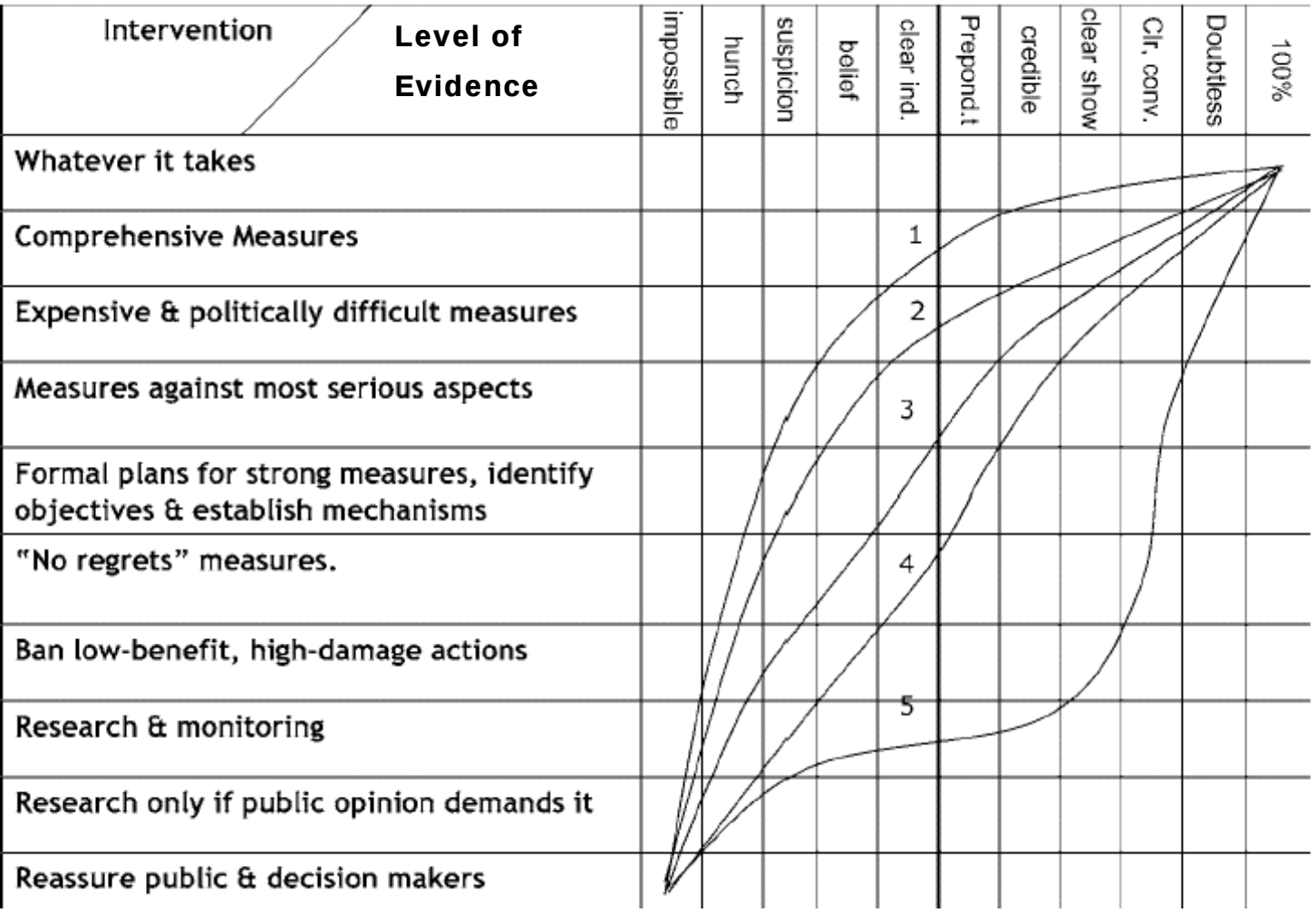
Reporting

- Make uncertainties explicit
- Assess robustness of results
- Discuss implications of uncertainty findings for different settings of burden of proof
- Relevance of results to the problem
- Progressive disclosure of information -> traceability and backing

Weiss 2003/2006 evidence scale

10. Virtually certain
9. Beyond a reasonable doubt
8. Clear and convincing Evidence
7. Clear showing
6. Substantial and credible evidence
5. Preponderance of the Evidence
4. Clear indication
3. Probable cause: reasonable grounds for belief
2. Reasonable, articulable grounds for suspicion
1. Hunch
0. No suspicion

Even where there is agreement on "level of evidence", there usually is substantial societal disagreement on what level of intervention is justified.



Attitudes according to Weiss 2003:

1. Environmental absolutist

2. Cautious environmentalist

3. Environmental centrist

4. Technological optimist

5. Scientific absolutist

Insights on uncertainty

- More research tends to increase uncertainty
  - reveals unforeseen complexities
  - Complex systems exhibit irreducible uncertainty (intrinsic or practically)
- Omitting uncertainty management can lead to scandals, crisis and loss of trust in science and institutions
- In many complex problems unquantifiable uncertainties dominate the quantifiable uncertainty
- High quality $\neq$ low uncertainty
- Quality relates to **fitness for function** (robustness, PP)
- Shift in focus needed from reducing uncertainty towards reflective methods to explicitly cope with uncertainty and quality