
cession Technologies and Evaluation

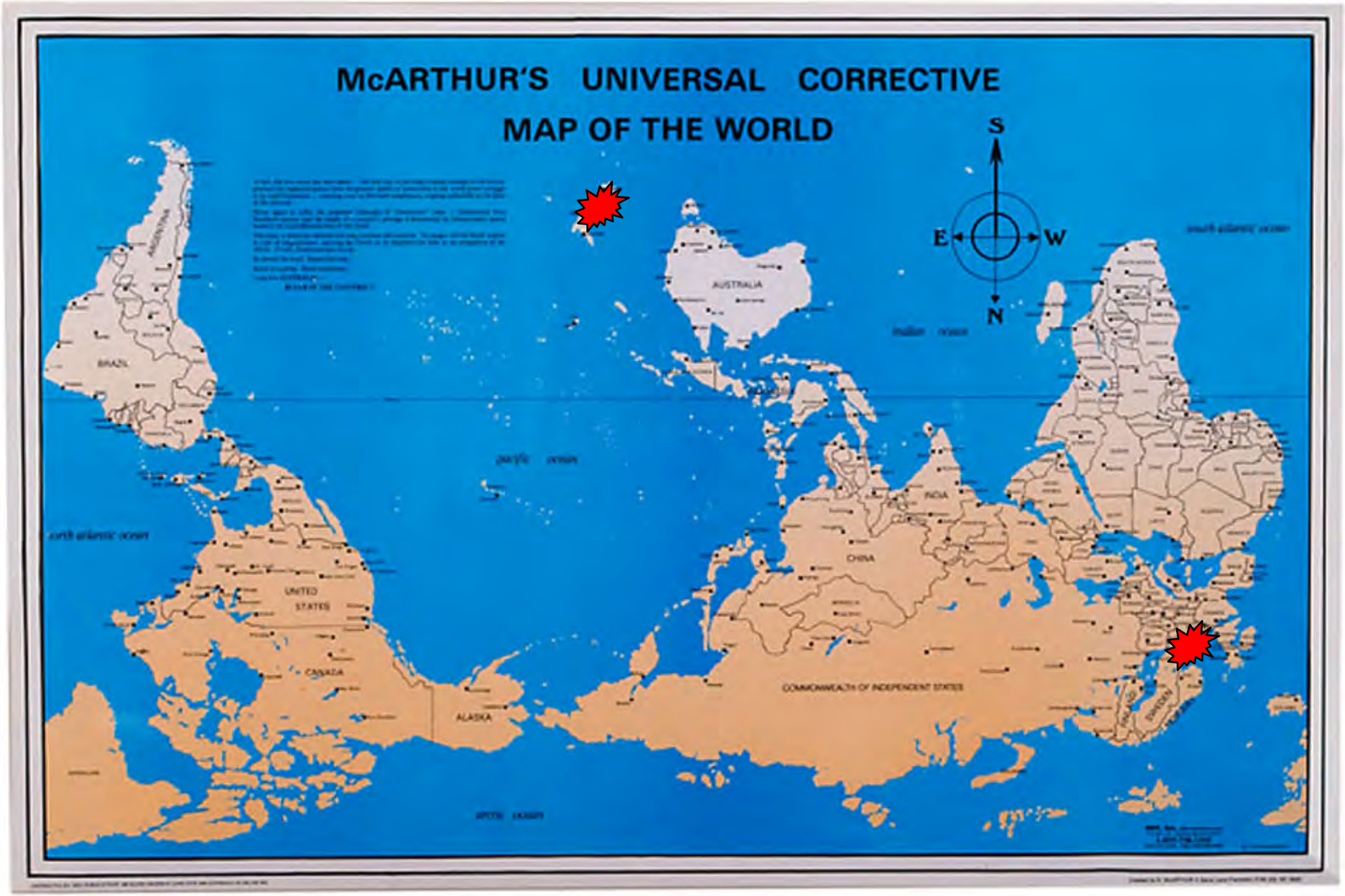
Decision Technologies and Evaluation

Jack Dowie
Professor of Health Impact Analysis
London School of Hygiene and Tropical Medicine

jack.dowie@lshtm.ac.uk

I am not young enough to know everything

Oscar Wilde


McARTHUR'S UNIVERSAL CORRECTIVE
MAP OF THE WORLD
S
E
W
N
AUSTRALIA
BRAZIL
INDIA
CHINA
UNITED STATES
CANADA
ALASKA
COMMONWEALTH OF INDEPENDENT STATES


lab test
p-value
guidelines
expert judgement
epidemiology
algorithm
bayes
consensus panel
randomised controlled trial
meta analysis
systematic review
reflective practice
health impact assessment
expert opinion
health economics
decision analysis
ethical code
health utility elicitation
case control study
evidence-based
cohort study
exposure-response model
policy committee
moral principle
cost effectiveness analysis
health impact analysis
QALYS
health impact estimation
autopsy
clinical judgement
patient preferences
case conference
hippocrates
risk assessment

5 key tensions

- Between Knowledge Technologies (KTs), Valuation Technologies (VTs) and Decision Technologies (DTs)
- Between more analytical and more intuitive modes of inquiry/understanding, valuation, and choice
- Between coherence-based and correspondence-based approaches to quality assessment
- Between Bayesian and non-Bayesian approaches to evidence synthesis (*inter alia*)
- Between absolutist and consequentialist approaches to values and ethics

Where does ‘evaluation’ fit in?

- Everywhere, depending on who is talking …and who is listening..
- An ‘evaluation’ might be
  - a KT for KT purposes
  - a KT for DT purposes
  - a VT for VT purposes
  - a VT for DT purposes
  - a DT for DT purposes
  - all or none of the above

Taxonomy of tasks and technologies

- To produce knowledge, elicit knowledge and evaluate knowledge claims we need a
  Knowledge Technology
- To establish what values are held and elicit value judgements we need a
  Valuation Technology
- To make decisions (including policies) we need a
  Decision Technology

DTs need inputs from KTs and VTs; transfer requires ITs and CTs

- To provide information we need an
  Information Technology
- To communicate we need a
  Communication Technology

Examples

- KT
  - ‘(clinical) opinion’, ..cohort study, RCT, lab test
- VT
  - ‘(clinical) judgement’, interview…, Standard Gamble, Time Trade-Off utility elicitation exercises
- DT
  - ‘(clinical judgement’), coin toss, meeting, …., pro-con checklist, … Decision Analysis
- IT/CT
  - nudge/wink, …., ppt presentation, Report with tables and graphs

A very inefficient KT
- Gary Larson

"What fits your busy schedule better, exercising one hour a day or being dead 24 hours a day?"

A very simple VT
– Randy Glasbergen

3 Risk factors

Black

Male

Baseball cap backwards

4 Mitigating factors

Loafers

Fed Ex envelope

Whistling Sondheim

Polo shirt

A moderately analytic DT

3 Risk factors

White

Male

Wrong neighbourhood

4 Mitigating factors

White

Groceries

Humming Motown

Over 40

– Garry Trudeau

BREAKTHROUGH IN BREAST DRUG TRIAL

By LORRAINE DAVIDSON

SCIENTISTS halted trials of a breast cancer drug – so more women could benefit from it immediately.

And doctors believe the breakthrough may help millions.

A total of 5,200 women from all over the world agreed to test Letrozole.

Half took it and half did not. All were cancer-free when the trial began.

Researchers found the drug cut the risk of the cancer recurring by 43 per cent and of spreading from one breast to another by 46 per cent.

The results were so dramatic the trial was halted less than halfway through so the group not on the drug could start taking it.

Breast cancer sufferers are usually put on drug Tamoxifen after diagnosis but tests show no benefits if it is taken for more than five years. Letrozole is the first drug aimed at stopping the recurrence of breast cancer – a big concern for former sufferers who have come off Tamoxifen.

Both treatments are for post-menopausal women.

Professor Ian Smith of The Royal Marsden Hospital in London where part of the trial was held said the results had "surpassed expectations".

Christine Fogg of Breast Cancer Care said: "The issue of recurrence is a great concern so the findings of this trial are encouraging." Two weeks ago it was revealed drug Anastrozole could also cut breast cancer risk.

KT ? DT

- **TRIAL CUT SHORT AFTER BENEFIT OF BREAST-CANCER DRUG IS PROVED**

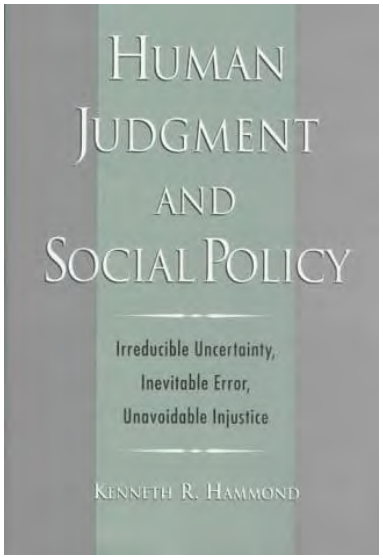
- AN INTERNATIONAL trial of a drug for breast cancer has been halted after early results showed it cut by 43 per cent the risk of the cancer returning in women already treated for the disease.

- Professor Ian Smith, head of the breast unit at the Royal Marsden Hospital, London, said: "This is one of the most important advances in the treatment of postmenopausal women with breast cancer, and is a further valuable step in preventing disease recurrence."

- But Professor Smith was among British specialists, including Professor Jack Cuzick, at Cancer Research UK, who criticised the decision to halt the trial early because the long-term effects of letrozole may now never be known.

Underlying framework of map

- ‘Cognitive Continuum’ framework suggests that any implementation of a KT, VT, DT, IT or CT can be
  - characterised by its Analysis-to-Intuition ratio / balance
  - ‘quality-assessed’ by its internal Coherence and its external Correspondence



Intuition

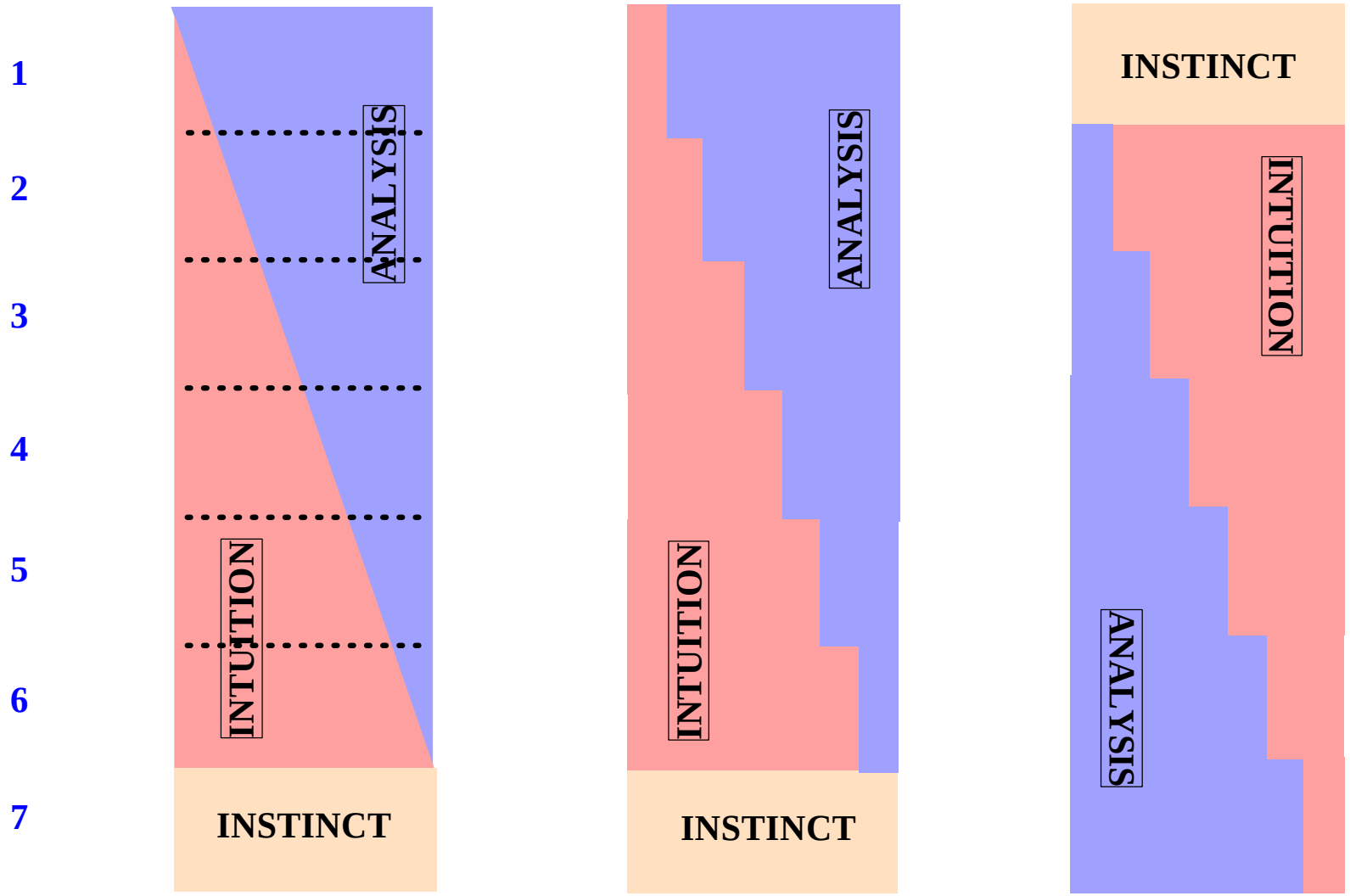
- rapid, unconscious processing of data
- combined by simple ‘averaging’ principle
- low consistency
- moderate accuracy - low potential for major errors??

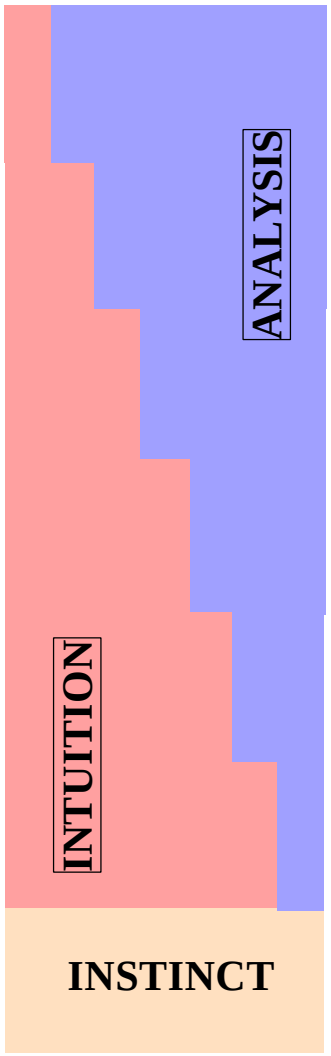
Analysis

- slow, conscious processing of data
- combined using more complex principles
- high consistency
- high accuracy - greater potential for major errors?? – but these will be more easily identifiable!

Intuition >Analysis

Analysis >Intuition


1
2
3
4
5
6
7
ANALYSIS
INTUITION
INSTINCT
ANALYSIS
INTUITION
INSTINCT
INSTINCT
INTUITION
ANALYSIS



Most precise and explicit

Definition of Concepts

e.g.
'health', 'risk'
'class'
'quality of life'

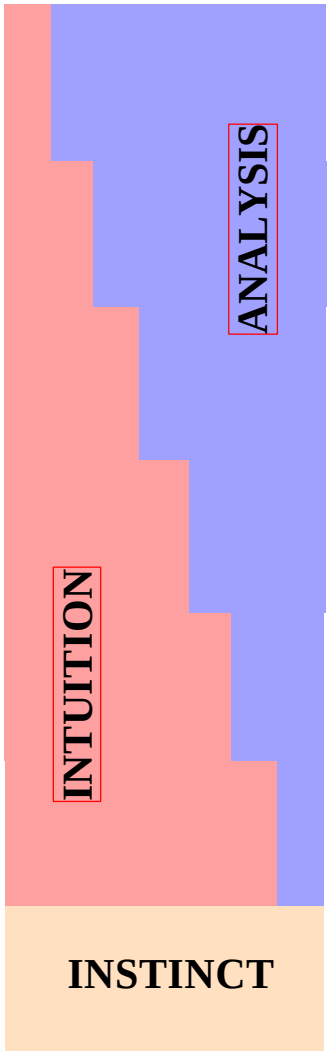
Specification of Relationships

e.g.
'health,
risk
and
class'

Measurement of Magnitudes

e.g.
'health'
'risk'
'quality
of life'

Least precise and explicit



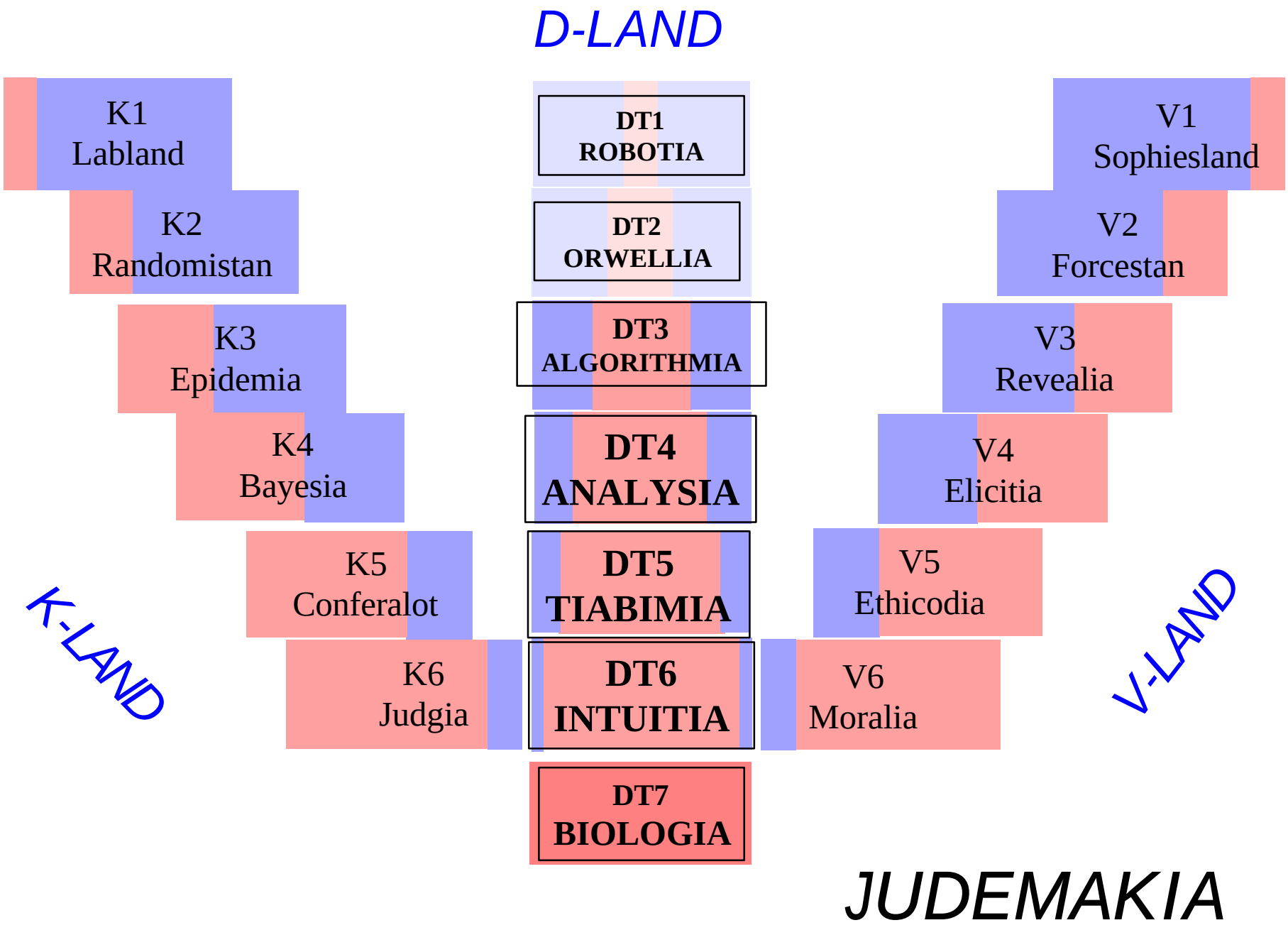
Highest demands and requirements

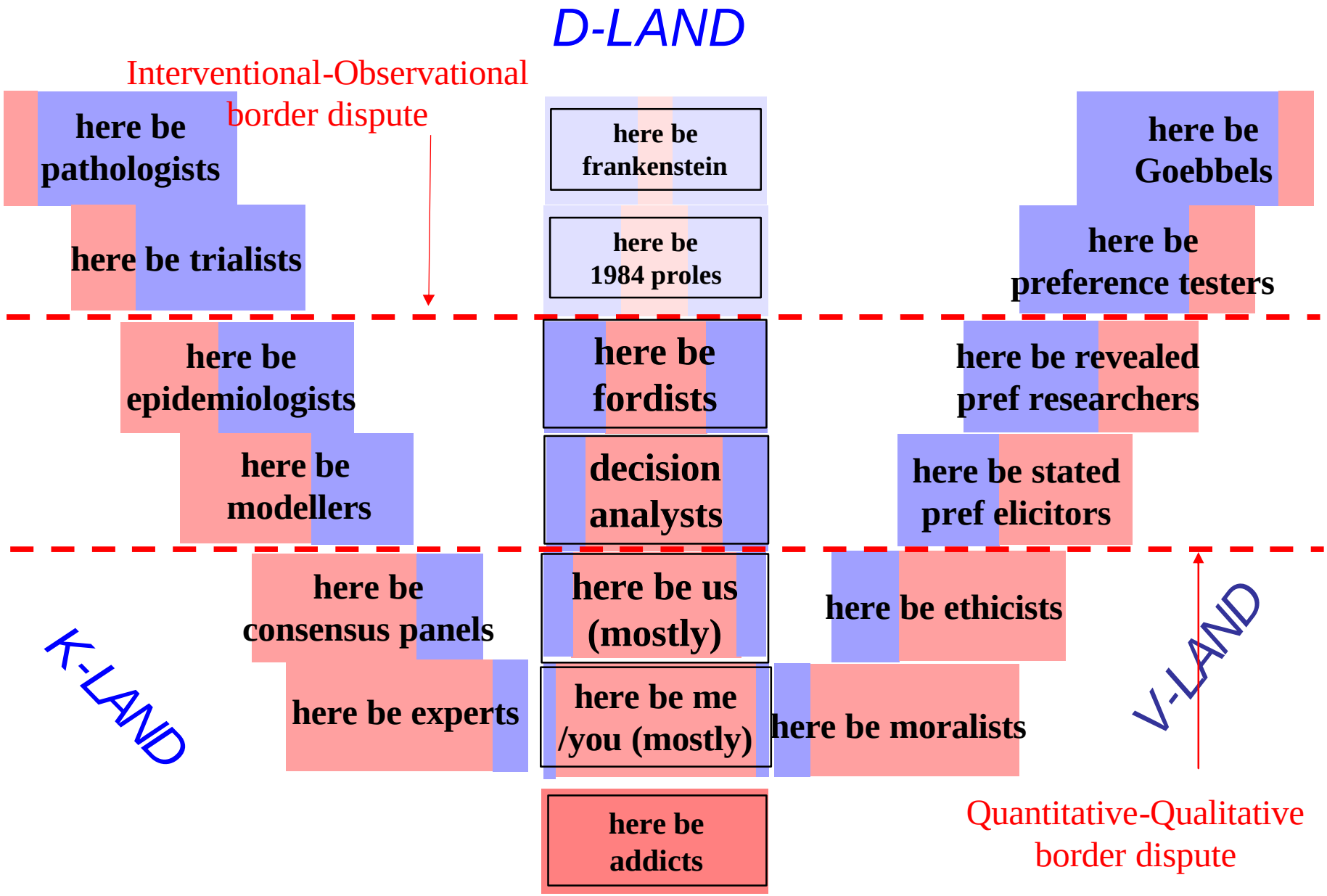
Control, ability to manipulate variables

Time and resources (??)

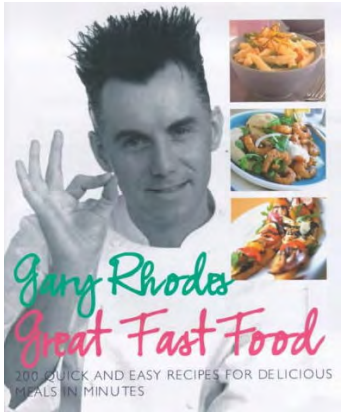
Analytical skills and capacities

Lowest demands and requirements


D-LAND
K1
Labland
K2
Randomistan
K3
Epidemia
K4
Bayesia
K5
Conferalot
K6
Judgia
K-LAND
DT1
ROBOTIA
DT2
ORWELLIA
DT3
ALGORITHMIA
DT4
ANALYSIA
DT5
TIABIMIA
DT6
INTUITIA
DT7
BIOLOGIA
V1
Sophiesland
V2
Forcestan
V3
Revealia
V4
Elicitia
V5
Ethicodia
V6
Moralia
V-LAND
JUDEMAKIA


D-LAND
Interventional-Observational
border dispute
here be pathologists
here be trialists
here be frankenstein
here be 1984 proles
here be Goebbels
here be preference testers
here be epidemiologists
here be modellers
here be fordists
decision analysts
here be revealed pref researchers
here be stated pref elicitors
here be consensus panels
here be experts
here be us (mostly)
here be me /you (mostly)
here be addicts
here be ethicists
here be moralists
K-LAND
V-LAND
Quantitative-Qualitative
border dispute


FRESH
MILK

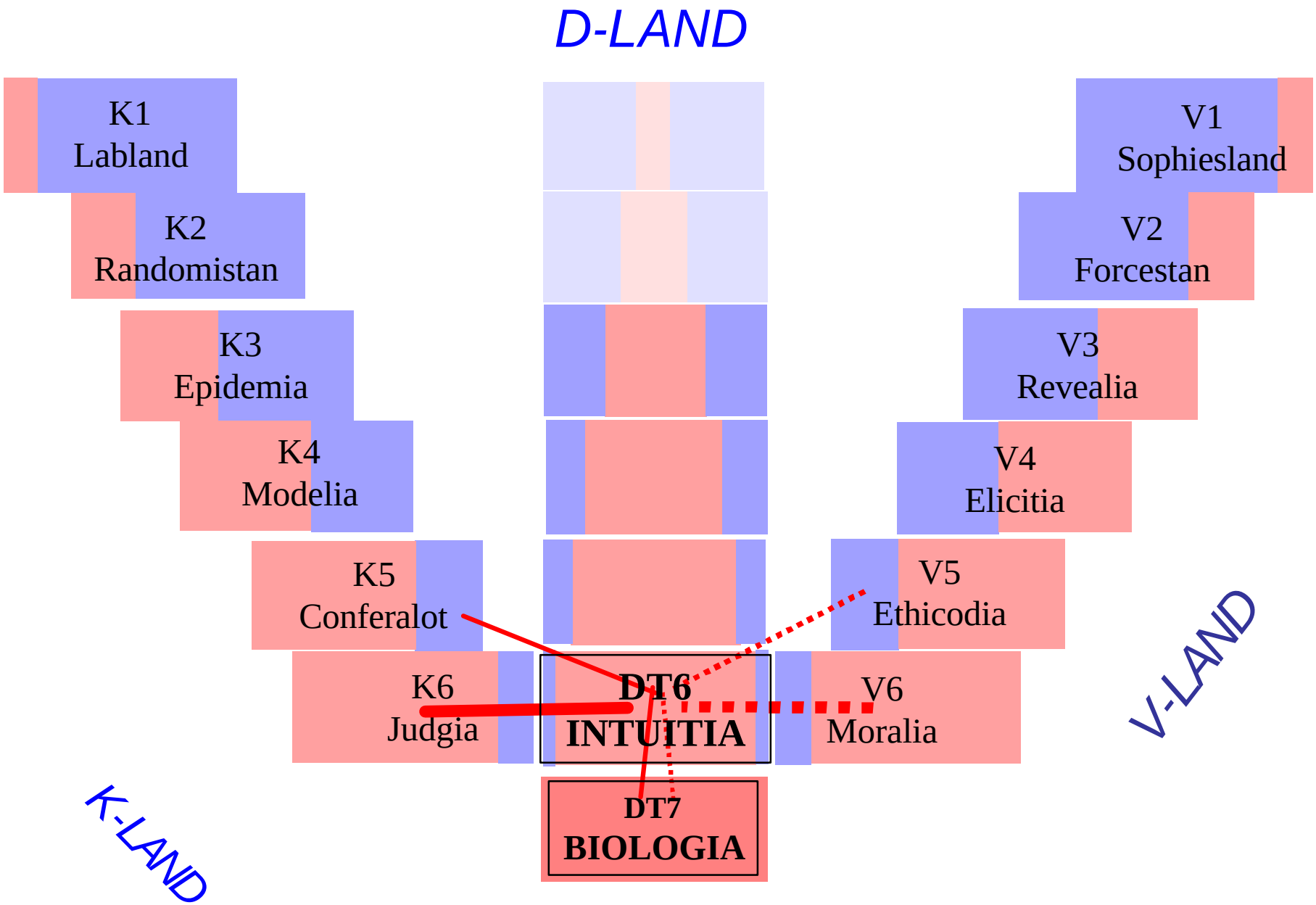

Gary Rhodes
Great Fast Food
200 QUICK AND EASY RECIPES FOR DELICIOUS MEALS IN MINUTES

The 4 basic Decision Technologies

- Biology (is Bayesian)
  - "I was compelled to do it, the emotional drive was so powerful"
  - "I acted instinctively, without thinking"
- Intuition (is Bayesian)
  - "I simply felt I could trust him/her"
  - "After all that experience I recognised the pattern instantly and knew what to do"
- TIABIM - Verbal reasoning (is Bayesian)
  - "We will Take Into Account and Bear in Mind all relevant considerations"
  - "I examined all the pros and cons in a balanced way"
- Decision Analysis (is Bayesian)

Instinct the only legitimate DT!

"So long as you only do what you honestly and instinctively believe is necessary in the heat of the moment, that would be the strongest evidence of you acting lawfully and in self-defence."



map3INTUITIA

Intuitive judgement as a DT/KT/VT

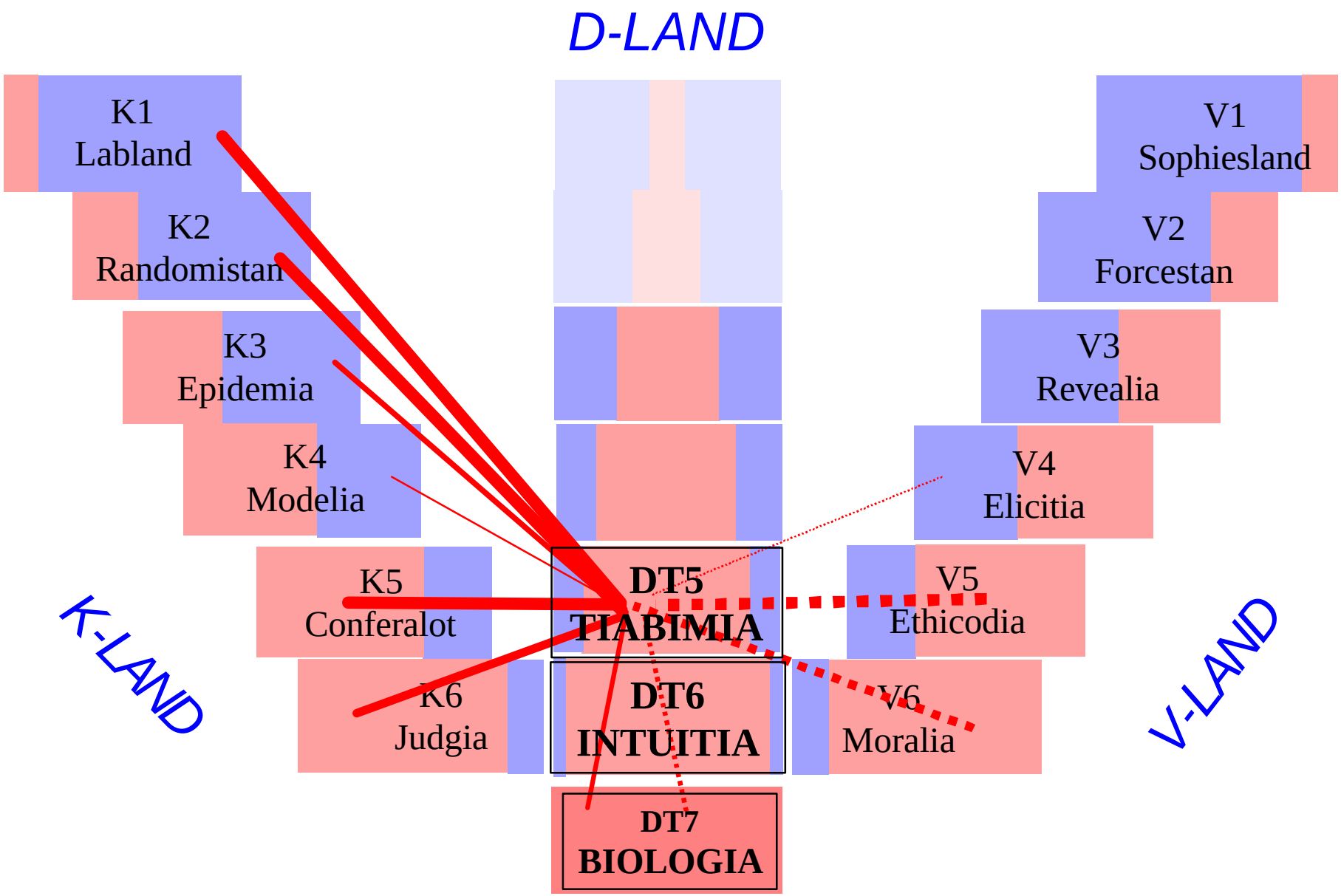
- Expertise
- Experience
- ‘Tacit knowledge’ (Michael Polanyi)
- ‘Reflective practice’
  - Donald Schön
  - Patricia Benner

- Evaluating ‘intuition’ – the vast problems
  - E.g. ‘ the waiter problem’ (‘treatment effect’)
    [OILS = Outcome Irrelevant Learning Situations]

‘Taking Into Account and Bearing In Mind’

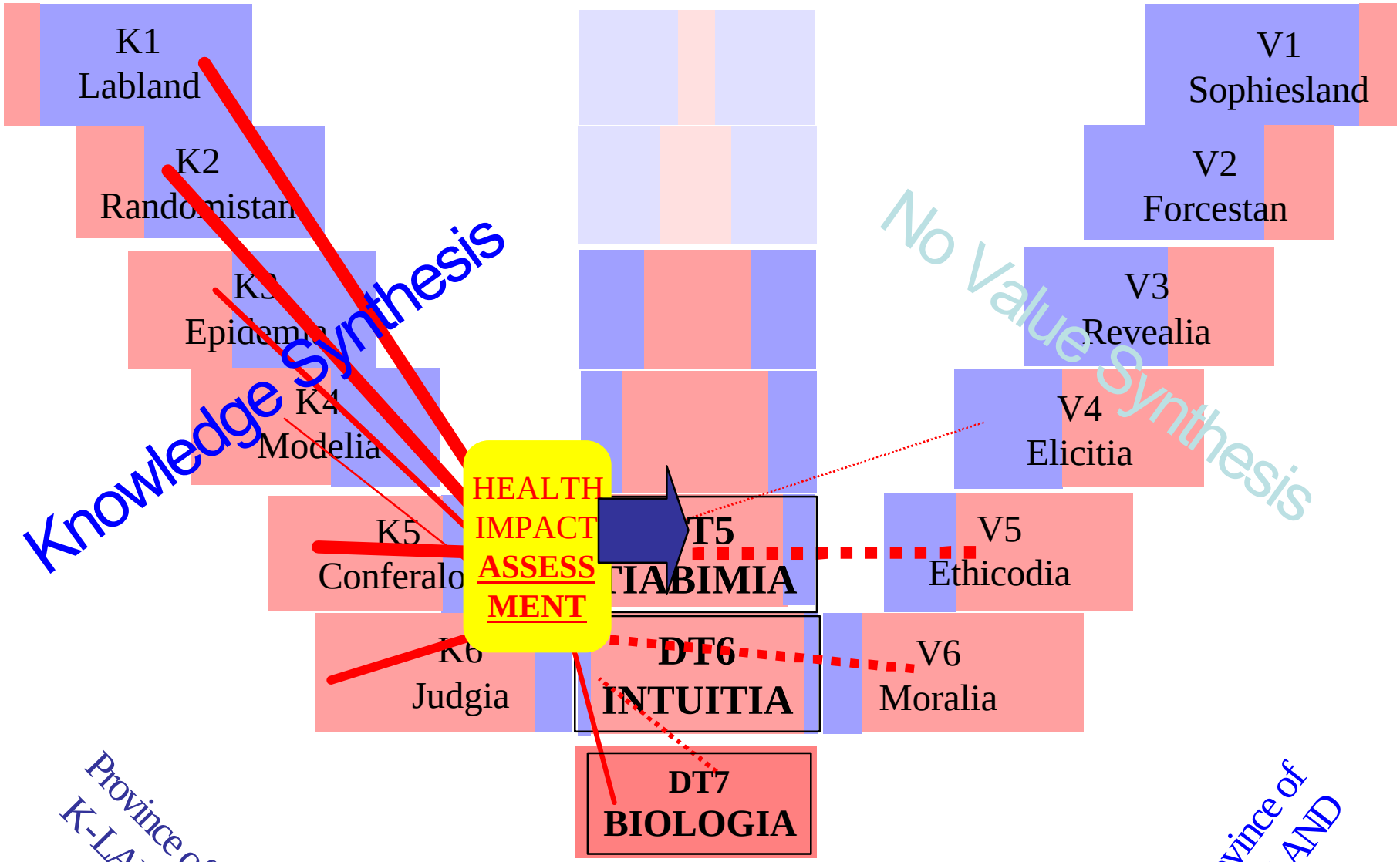
- ‘Taking things into *account*’
- ‘Giving considerations due *weight*’
- ‘Establishing the right *balance’*
- ‘Keeping things in *proportion’*
- ‘Taking a *measured* view’
- ‘Bringing everything into the *equation’*
- ‘*Figuring* it out’
- ‘Seeking a *degree* of consensus’
- ‘*Gauging* the impact’
- ‘Making sure things *add* up’

  - but TIABIM DT is basically *qualitative* discourses


D-LAND
K1
Labland
K2
Randomistan
K3
Epidemia
K4
Modelia
K5
Conferalot
K6
Judgia
V1
Sophiesland
V2
Forcestan
V3
Revealia
V4
Elicitia
V5
Ethicodia
V6
Moralia
DT5
TIABIMIA
DT6
INTUITIA
DT7
BIOLOGIA
K-LAND
V-LAND

TIABIM – the verdicts

- TIABIM is fine, no problems, flaws
- TIABIM is not perfect, it does produce lots of bad/poor decisions, but this is because
  - the wrong people with the wrong values dominate – bring in the true/right ‘stakeholders’ and it will be fine
  - we have the right people with the right values, but they lack knowledge/information/evidence – supply them with better k/i/e and it will be fine
  - we have the right people but our TIABIM processes need improvement
  - we have the right people but many currently lack the relevant TIABIM skills - build their capacity in these and it will be fine



Republic of D-LAND

map5HIAss

Why is TIABIM (and HIAss) so popular?

- TIABIM doesn't require clear and explicit separation of knowledge and values
- TIABIM doesn't require explicit and precise statement of anyone's knowledge/beliefs (in form of probabilities)
- TIABIM doesn't require clear explicit and precise statement of anyone's values (as quantified preferences)

E-BPH and HIAss?

- ‘Evidence-Based Public Health’
- ‘Health Impact Assessment’
- Both, in their different ways, are attempts to address the acknowledged weaknesses of Bayesian Decision Intuition (known as ‘clinical judgement’ and ‘public health judgement’ respectively)
- But without moving to Bayesian Decision Analysis or accepting consciously that decision making is *necessarily* Bayesian

There is no alternative to Bayes

- There is no such thing as a *non-Bayesian* approach to decision making, whether it is about health, health inequalities or anything else.
- If one is deciding under uncertainty – and genuine decisions are always made under uncertainty - one is, *at the moment of decision*, being a Bayesian in the key sense of that term. … assessing the probabilities of *unique* events or states (not their long run frequency under alternative hypotheses as in mainstream statistics)
- if one is making decisions (not inferences) then one must be assessing the probabilities of *unique* events or states and the only alternatives are some form of Bayesian Decision Intuition (BDI) or Bayesian Decision Analysis (BDA), not some non-existent, non-Bayesian alternative

Why are non-Bayesians?

- Why are non-Bayesians not prepared to produce and offer the probabilities needed by decision makers?
- Refusal means that decision makers are required to
  - perform covert and/or unwitting transformation of what they are offered into what they need (when they engage in BDI), or
  - required to carry out explicit and open transformational work (when they engage in BDA)?
- Two aspects to what is, at root, a single reason
  - clashing concepts of what a probability is
  - the use of prior beliefs in assessing the significance of a new piece of evidence

The nature of probability

- For Bayesians (in the words of Bruno de Finetti) there is no such *thing* as probability. Probabilities are properties of individual minds – a probability for something is simply an individual's uncertainty about it quantified as a degree of belief between 0 and 1.0.
- Probabilities are not properties of the external world, physical, biological or social. They are not the long run frequencies of events as the non-Bayesian 'frequentist' would have it. Accordingly a Bayesian will (if not being lazy or over-polite) always talk of '*my/his/her* probability *for*' an event happening (it raining tomorrow) or a state arising (me getting wet) never '*the* probability *of*' it.
- Frequencies undoubtedly can be discerned in, or imposed on the real world, and most Bayesians are committed to frequencies as a very useful basis for arriving at their degrees of belief. But they point out that in order to calculate a particular frequency one must *construct*, via 'subjective judgments', the observations to include in the frequency calculations (both numerator and denominator).

Prior beliefs and Bayes Theorem

- Bayes theorem states that the **P**robability of **A**buse **given** **B**ruises (what we usually want to know for decision making purposes) is equal to the **P**robability of **B**ruises **given** **A**buse multiplied by the **P**robability of **A**buse, the result divided by the **P**robability of **B**ruises.
- Extracting the emboldened capitals and using a vertical bar for 'given' gives us the standard simple form of Bayes theorem:
- **P(A|B) = [P(B|A) x P(A)] / P(B)**
- Our probability for an event/hypothesis (Abuse) after we acquire a new piece of evidence (Bruises) is correctly calculated by multiplying our probability for it before we received the new evidence by the probability of receiving this evidence given that our hypothesis is true and dividing the result by the probability of receiving this evidence irrespective of whether our hypothesis is true or false.

Q

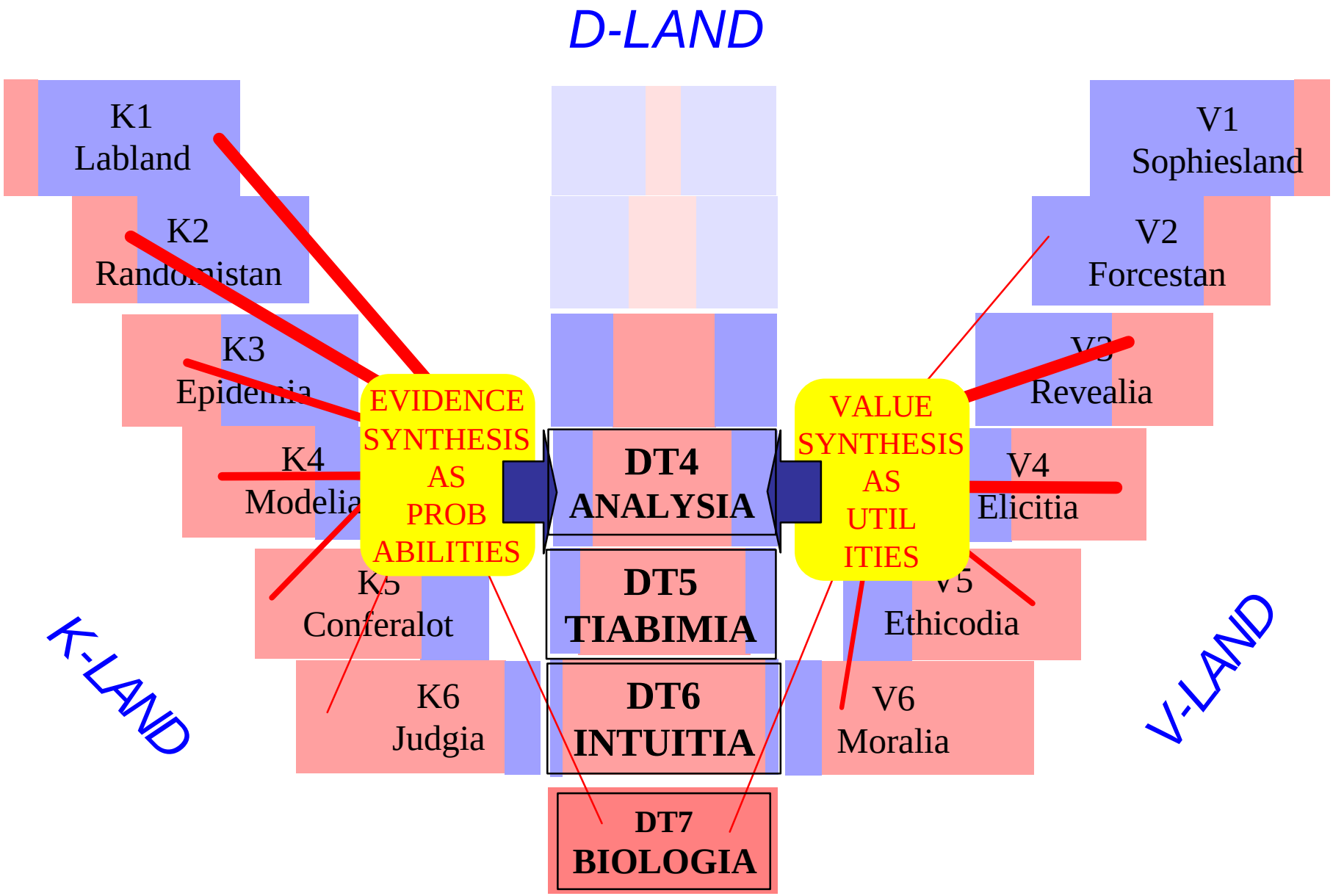
- A person has just tested positive for a condition (Cancer) which has a prevalence of 1% in the community.
- The test is a good one in the sense that it has a Sensitivity (True Positive Rate) of 90% and a Specificity (True Negative Rate) of 80%.
- What is the chance that the patient has Cancer?

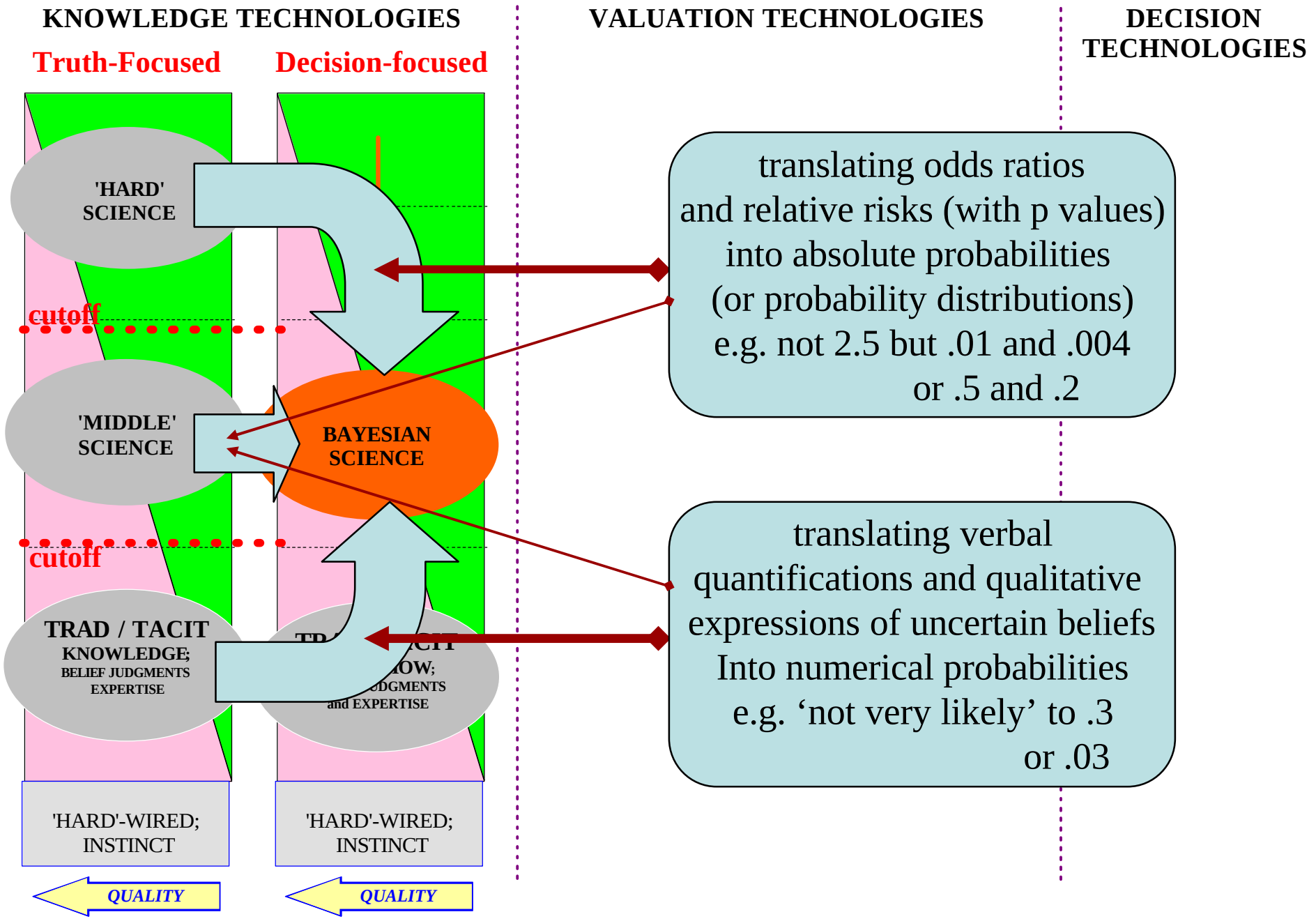
A

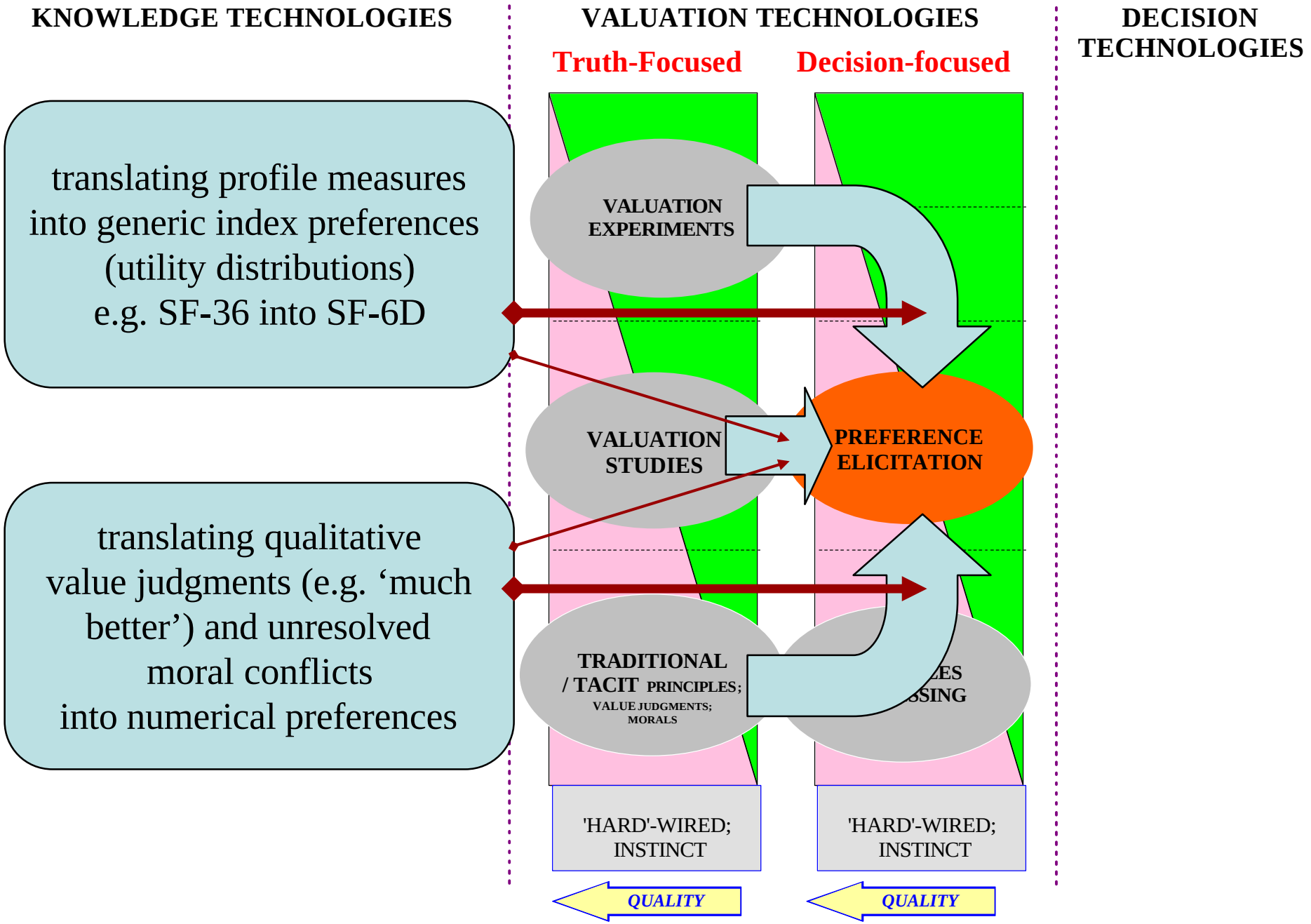
- The intuitive responses (over a wide range of audiences and countries) cluster in the 80-90% range, reflecting the TPR and TNR.
- The correct answer is 4.3%.
- The 20% False Positive Rate (1 – Specificity) applies to the 99% of the population who do not have Cancer and results in 198 False Positives compared with only 9 True ones.
- Of the 207 total positives only 9, or 4.3%, will be TP and this is therefore the Predictive Value of a positive test result

The irrelevance of inference

- Most Decision Analysts would see little point in being Bayesian if they were not Decision Analysts, while fully accepting that in order to be Decision Analysts they must be Bayesians.
- The key difference is not between Bayesian and non-Bayesian approaches to *statistical inference* but between the Decision Analytic and conventional approaches to *decision making*; conventional approaches that, whatever their surface appearance, can all be characterised as some form of BDI (either simple Intuition or TIABIM) .
- The statistical conflict is only important because it is a major contributor to the quality problems of BDI, largely through the cognitive burden it imposes – or should impose - on decision makers as they struggle with the impossible task of transforming the decision-irrelevant format of scientific output into what they need.
- A cognitive burden typically reduced by the use of inappropriate heuristics and unwitting misinterpretations.


D-LAND
K1
Labland
K2
Randomistan
K3
Epidemia
K4
Modelia
K5
Conferalot
K6
Judgia
K-LAND
EVIDENCE
SYNTHESIS
AS
PROB
ABILITIES
DT4
ANALYSIA
DT5
TIABIMIA
DT6
INTUITIA
DT7
BIOLOGIA
VALUE
SYNTHESIS
AS
UTIL
ITIES
V1
Sophiesland
V2
Forcestan
V3
Revealia
V4
Elicitia
V5
Ethicodia
V6
Moralia
V-LAND


KNOWLEDGE TECHNOLOGIES
VALUATION TECHNOLOGIES
DECISION TECHNOLOGIES
Truth-Focused
Decision-focused
'HARD' SCIENCE
cutoff
'MIDDLE' SCIENCE
BAYESIAN SCIENCE
cutoff
TRAD / TACIT KNOWLEDGE; BELIEF JUDGMENTS EXPERTISE
and EXPERTISE
'HARD'-WIRED; INSTINCT
'HARD'-WIRED; INSTINCT
QUALITY
QUALITY
translating odds ratios and relative risks (with p values) into absolute probabilities (or probability distributions) e.g. not 2.5 but .01 and .004 or .5 and .2
translating verbal quantifications and qualitative expressions of uncertain beliefs Into numerical probabilities e.g. 'not very likely' to .3 or .03

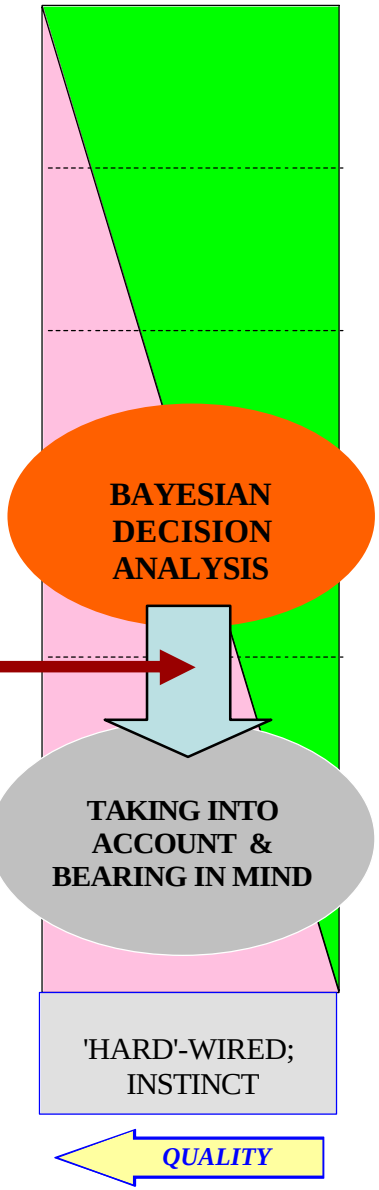

KNOWLEDGE TECHNOLOGIES
VALUATION TECHNOLOGIES
Truth-Focused
Decision-focused
DECISION TECHNOLOGIES
translating profile measures into generic index preferences (utility distributions) e.g. SF-36 into SF-6D
translating qualitative value judgments (e.g. 'much better') and unresolved moral conflicts into numerical preferences
VALUATION EXPERIMENTS
VALUATION STUDIES
TRADITIONAL / TACIT PRINCIPLES; VALUE JUDGMENTS; MORALS
PREFERENCE ELICITATION
'HARD'-WIRED; INSTINCT
'HARD'-WIRED; INSTINCT
QUALITY
QUALITY

KNOWLEDGE TECHNOLOGIES

VALUATION TECHNOLOGIES

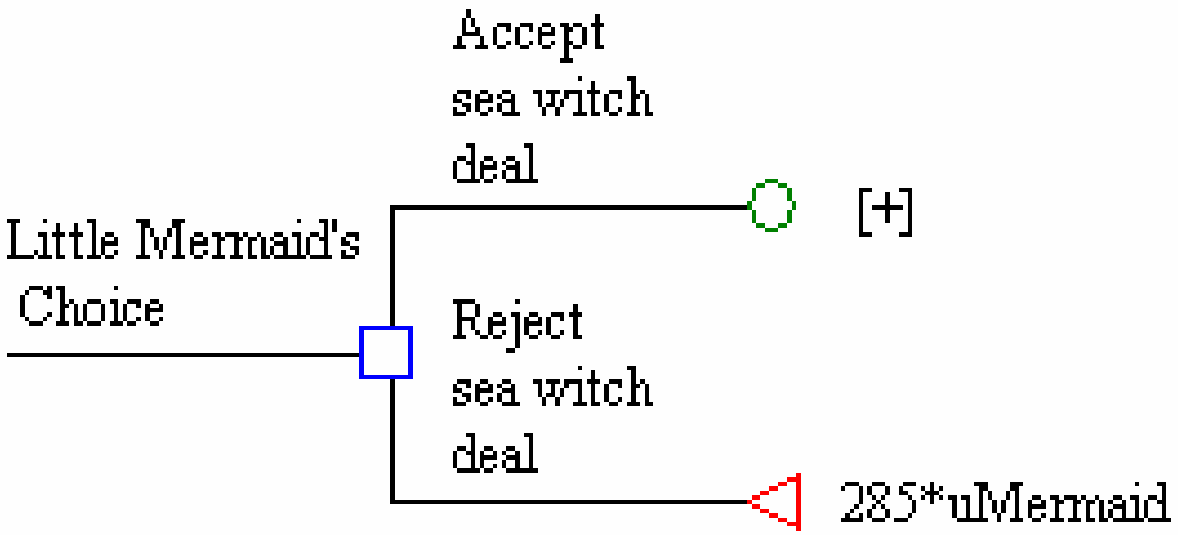
DECISION TECHNOLOGIES

attempts to TIA results of BDA without any means of ensuring coherence or transparency and no guarantee decision is of better net quality (except by assumption)



Den Lille Havfrues Beslutning
The Little Mermaid's Decision

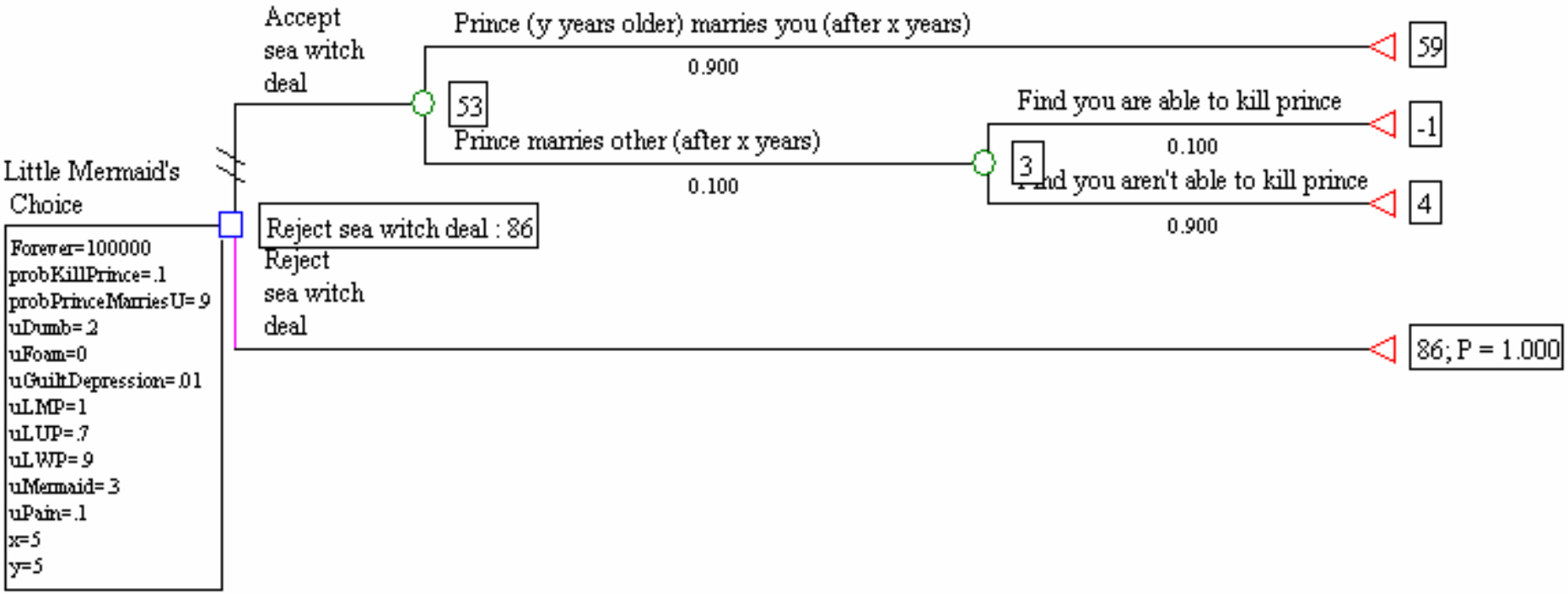
Acceptere Havheksens handel



Afslå Havheksens handel

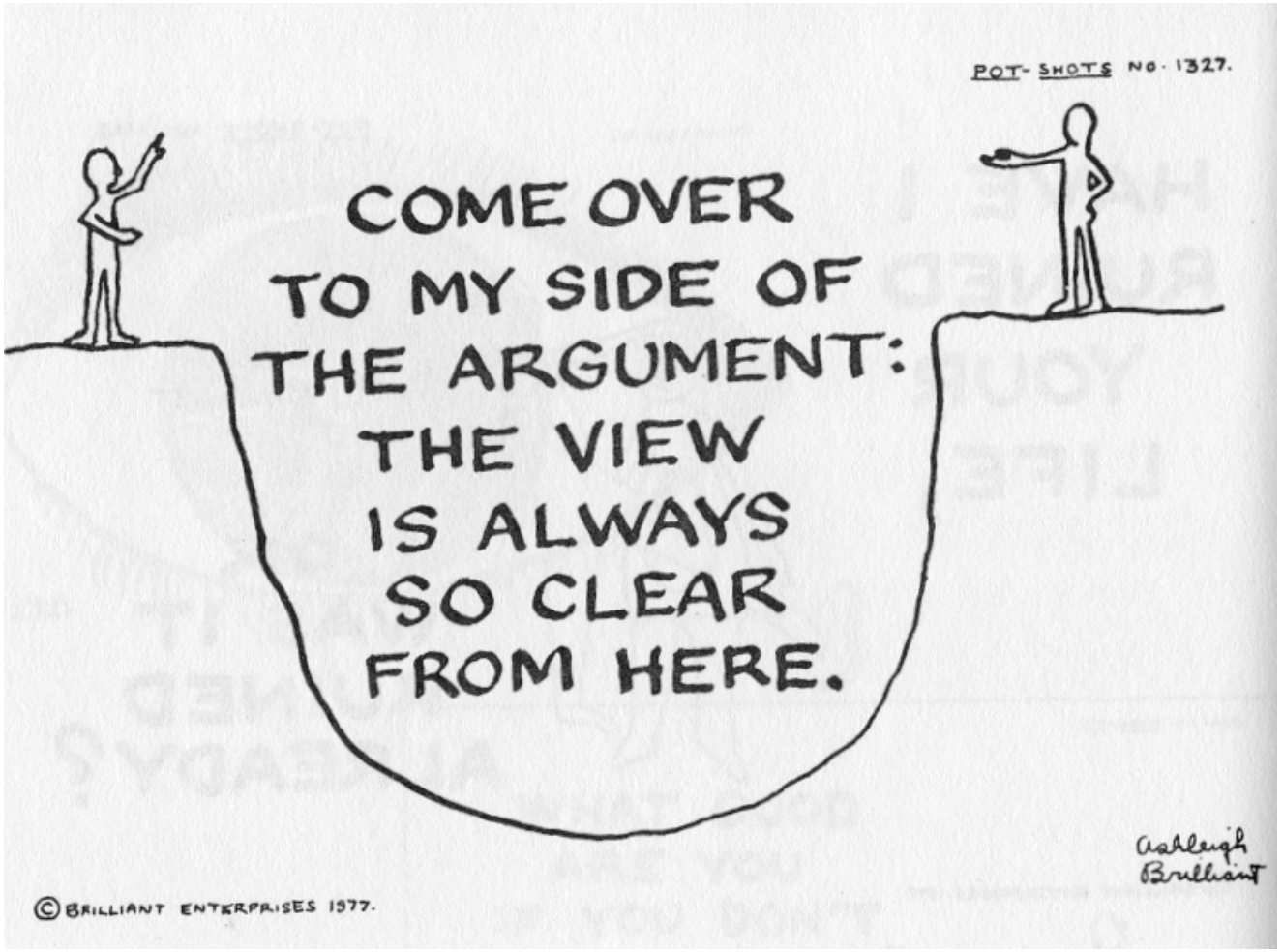
Quality of Life in various states

State	QOL before deductions	QOL after deductions
Living Married to Prince	**1**	**.7**
Living as Widow of Prince	**.9**	**.6**
Living in castle Unmarried to Prince	**.7**	**.4**
Mermaid (no deductions)	**.3**	**.3**
Mermaid (after killing prince)		**.25**
Foam	**0**	
QOL Deductions for:		
Loss of voice	**.2**	
Pain on every movement	**.1**	
Guilt/depression from murder	**.05**	


Little Mermaid's Choice
Forever=100000
probKillPrince=.1
probPrinceMarriesU=.9
uDumb=.2
uFoam=0
uGuiltDepression=.01
uLMP=1
uLUP=.7
uLWP=.9
uMermaid=3
uPain=.1
x=5
y=5
Accept sea witch deal
Reject sea witch deal
Reject sea witch deal : 86
53
Prince (y years older) marries you (after x years)
0.900
59
Prince marries other (after x years)
0.100
3
Find you are able to kill prince
0.100
-1
Find you aren't able to kill prince
0.900
4
86; P = 1.000

The inappropriateness of ‘scientific’ cut-offs

- Decisions always requires integration of values and uncertainties
- Policy making should always be seen (and referred to) as a ‘*value-based and science-informed* activity’, since it can never be a ‘science-*based* activity’
- Absence of high / ‘acceptable’ quality evidence (by *scientific* criteria appropriate for a truth-focused KT) doesn’t mean that decisions can be postponed and resort to evidence of lower / ‘unacceptable’ quality avoided (the only question is whether resort is made are implicit or transparently and explicitly)
- Setting data quality cut-offs in accordance with the evaluation criteria appropriate in science (typical in Evidence-based Practice checklists – and NICE) is therefore inappropriate for policy/decision making, where the Best available evidence has to be used.


POT-SHOTS NO. 1327.
COME OVER TO MY SIDE OF THE ARGUMENT: THE VIEW IS ALWAYS SO CLEAR FROM HERE.
Ashleigh Brilliant
©BRILLIANT ENTERPRISES 1977.