



Presenter: Alistair

Moderator: Carol

A few thoughts on the nature of
Statistical Significance
&
Operational Significance



Topic 6 : Statistical and Operational Significance



Action Plan 5 Workshop

March 2002

Topics:

The relationship between the two types of significance

The meaning of results

The needs of different stakeholders

Presenting results in responsible ways/ the role of validation and validators



Statistical Significance:

- ✓ Is telling us about the probability that an observed difference between two distributions is due to chance or not

It is not telling us about:

- ✗ the size of a difference
- ✗ the replicability of a difference
- ✗ the operational significance of a difference



Operational Significance

- ☒ may provide an explanation of behaviour in the real world system
- ☒ risks to have a negative impact on the real world system
- ☒ may be used as a basis for control (effecting predictable change in a real world system)

The Context:

- ? A statistically significant factor may not be operationally significant
- ? A factor may be operationally significant without being statistically significant



For an effect to be operationally USEFUL

- ☒ it should be measurable
- ☒ it should be reproducible
- ☒ more it should be predictable in its magnitude as well as its occurrence
- ☒ we should have an understanding of the process underlying the effect and be able to explain (and manipulate) it



Operational Significance and the HITL Experiment

- ⊗ basic trade-off, specific answers versus generally applicable conclusions
- ⊗ face validity, test-retest validity
- ⊗ how often do we repeat a study or test transfer?

How do we deal with differences between user preferences and observed performance?

- ⊗ are our measures adequate?
- ⊗ do controllers like tools that don't really help? why?
- ⊗ are we asking the wrong questions? Demonstrating our lack of understanding?



Stakeholder Information Needs

Validation:

The process through which a desired level of confidence in the ability of a deliverable to to operate in a real-life environment may be demonstrated against a pre-defined level of functionality, operability and performance.

The Context:

A community of stakeholders with different interests, perspectives and criteria. These stakeholders need to be identified and their needs understood if **consensus** is to be built.



The Probable Consequence

Some stakeholders will require 'scientific' evidence.
Others will have different needs,

- e.g. experiential (hands-on)
- assessment of the quality of a process
- perception of involvement
- trust in the agents of change



To Build the validation consensus

We must furnish the appropriate evidence and communicate in an appropriate form

BUT

It should be the best information we can obtain.

(This may mean 'scientific' collection even for the 'non-scientific' presentation).

Distinction between the way that data is collected and analysed and the way it is presented?



Roles and Responsibilities of Validation Practitioner's

- ⊗ as honest broker's between the different stakeholder agendas
- ⊗ as the people who are prepared to 'say no'?
- ⊗ as those who question assumptions (especially their own)?
- ⊗ Willingness to accept the responsibility?
- ⊗ Institutional structure which supports such a role for validation - not just lip service?