Issues related to quantitative methods and data analysis in PER

David E. Meltzer
Iowa State University
Richard R. Hake
Indiana University

Issues Related to Quantitative Methods and Data Analysis in PER

David E. Meltzer *
Dept. of Physics and Astronomy, Iowa State University, <dem@iastate>
&
Richard R. Hake †
Physics Dept., Indiana University, <rrhake@earthlink.net>

---

* Partially supported by NSF Grant DUE-9981140 & REC-0206683.

† Partially supported by NSF Grant DUE/MDR-9253965.

Material drawn from:

(a) Hake (2001) " Suggestions for Administering and Reporting Pre/Post Diagnostic Tests," online as ref. 14 at < **http://www.physics.indiana.edu/~hake** >;

(b) Hake (2002a) " Lessons from the Physics Education Reform Effort," online at < **http://www.consecol.org/vol5/iss2/art28** >.

Methodological Issues that are **Always** Relevant

(*implicitly or explicitly)*

- **Validity**
- **Reliability**
- **Statistical Significance**
- **Pedagogical Significance**
- **Representativeness of sample**
- **Reproducibility**

Validity

- Does your instrument provide data that could actually answer your research question?
  - Sample Flaw: To answer question, “Do students understand concept *A*?,” give them problem requiring knowledge of concepts *A, B, C,* and *D.*
    - Previous knowledge of specific mathematical tools or formal representations may need to be considered

Reliability

- Is instrument internally consistent, that is, do different components of instrument measure (more or less) the same property?
- If you made the same measurement again (all conditions apparently identical), would your instrument yield the same result?
- Would minor variations in your test items (e.g., slight contextual or representational changes) lead to large variations in results?

Statistical Significance

- Is there a substantial probability ($\geq \approx 10\%$) that your result might have occurred purely by chance?
  - Do you have a measure of variance, or can one be estimated?
  - Can standard statistical tests be applied?
    - Parametric [e.g., linear correlation] and non-parametric [e.g., ranking]

Pedagogical Significance

- Is the observed effect likely to be of practical significance in classroom?
- Are there cost-benefit relationships implied in the magnitude of the observed effect?

Representativeness of Sample

- Is your sample representative of larger group from which it is (implicitly or explicitly) drawn?
- Are samples from different groups being compared equivalent in all respects except variable being investigated?
  - What are possible relevant variables?
  - How have you controlled these variables?
  - Have you done random selection? If not, what alternatives were used?

Reproducibility

- Did you repeat experiment?
- Did anybody else repeat the experiment?
- Are your results substantially different from what others have observed, or are they otherwise very surprising?
  - If so, better check again!