

## REVIEWERS' COMMENTS

- [Comment on Clinical vs Statistical Significance](#): Alan Batterham
- [Comment on Qualitative vs Quantitative Research Designs](#): Keith Davids

[Reprint pdf](#) · [Reprint doc](#)

**Comment on Clinical vs Statistical Significance**

Alan M Batterham, Department of Sport and Exercise Science, University of Bath, Bath BA2 7AY, United Kingdom. [Email](mailto:alan.batterham@bath.ac.uk). [sportsci.org/jour/0103/inbrief\\_comments.htm#clinical](http://sportsci.org/jour/0103/inbrief_comments.htm#clinical), 2001 (461 words)

This item is a very useful short critique of the poor scientific practice of over-reliance on tests against the null hypothesis with arbitrary P values. I hope that authors and reviewers take note. I wondered if some of the key points could be highlighted with a pertinent quote or two to hammer home the message? For example, the oft-quoted "Surely God loves the 0.06 nearly as much as the 0.05?" (Rosnow and Rosenthal, 1989) would illustrate well the point that one's article is much more likely to be accepted if  $p < 0.05$ . I have found that such quotes really help the lay reader and statistically naive to grasp the point.

The points regarding a confidence interval approach to estimation are well made and will help get across the key question in research: is the effect big enough to be scientifically/clinically/practically relevant/important? I like the comments regarding the potential problems with using 95% confidence limits. This echoes your critique of the Bland/Altman 95% limits of agreement for quantification of reliability and the superiority of the typical error. Incidentally, Sterne and Smith (2001) have also opted for 90% confidence limits, but they did not overtly provide a justification for their recommendation. My reading of it is that they were proposing it more as a deterrent to the practice of using the 95% limits as a surrogate means of testing against the null hypothesis at the 0.05 alpha level and thus falling into the same trap. I agree that choice of limits less than 95% may help overcome the firmly entrenched 0.05 alpha level.

Hopefully, your article will help people to put the research question ahead of the straw man of the null hypothesis and thus not allow the statistics to detract from the ultimate vehicle for making inferences—the data themselves. I think the key obstacle to this process is that, unlike a yes/no decision based on some arbitrary alpha level, it requires genuine thought and intellectual rigor to determine the smallest worthwhile effect for the variable in question! Yet, that is, or should be, the crux of our scientific endeavor.

Finally, I wondered if the points in the last paragraph could be stated even more emphatically or combatively? The misconceptions about what null hypothesis testing does and does not tell us are close to universal and will not be overcome without radical and persistent challenge!

[Back to In Brief](#)

Rosnow RL, Rosenthal R (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist* 44, 1276-1284

Sterne JAC, Smith GD (2001). Sifting the evidence—what's wrong with significance tests. *BMJ—British Medical Journal* 322, 226-231

## Comment on Qualitative vs Quantitative Research Designs

Keith Davids, Department of Exercise and Sport Science, Manchester Metropolitan University, Alsager, Cheshire ST7 2HL, UK. [Email](mailto:Email). [Sportsjournal 5\(3\), sportsjournal.org/jour/0103/inbrief\\_comments.htm#qual](http://sportsjournal.org/jour/0103/inbrief_comments.htm#qual), 2001 (585 words)

It is good for scientists to have a range of methodological approaches to tackle the large variety of experimental and practical questions in sport. Practical work in coaching of team sports is often biased towards quantitative analysis of the group, whereas historically psychotherapists and movement rehabilitation therapists have preferred to treat each patient on an individual basis. Single case studies are still relatively rare in the sports sciences, although they are more common in the behavioral sciences (Schöllhorn & Bauer, 1997). They are particularly useful in researching performance of elite able-bodied and disabled athletes who are available only in small numbers.

The emphasis on single subject designs and case studies recognizes the significant amount of variability in human behavior. For example, in sport science one aim of group work has been to identify key commonalities in movement patterns that can act as a reference point for learners in skill acquisition. During the modeling process in skill acquisition, these reference values can take the form of an optimal kinematic pattern for learners. Problems with the group approach can arise if the reference values for a common optimal pattern for all learners are based on the performance of one individual, for example a skilled athlete. The weaknesses with this approach are based on the well-documented problems of providing average or summarized feedback to groups of learners. Due to the unique constraints on each individual learner, it is likely that group-based feedback will provide a large proportion of irrelevant information for each individual. Traditional group-based analyses tell us only part of the story, as each individual attempts to find their own solutions to typical movement problems.

A dynamical systems perspective on movement coordination and control encourages a case study approach by treating each individual performer as a unique system learning to interact with the environment. Newell, Liu and Meyer-Kress (2001) have recently shown how the traditional approach of averaging data for groups across conditions and trial blocks may have masked the presence of different types of learning curves in individuals (e.g., exponential, S-shaped, sudden discontinuous), and they have questioned the ubiquity of the power law of (motor) learning. Averaging scores over individual participants and trial blocks ignores the fact that laws of learning should reflect both transitory and persistent changes in behavior over time, whereas the power law approach treats transitory effects as random-like behavior that can mask the persistent trend. Statistical techniques of pooling group data or blocking trials have an important role to play in quantitative research methods, in order to examine central tendencies and dispersion, but they may limit insights into the way that individuals solve coordination problems. For this reason, new case-study methodologies such as coordination profiling (Button and Davids, 1999) and self-organizing maps (Bauer and Schöllhorn, 1997) are emerging in motor behavior research to determine how each individual solves coordination problems.

### [Back to In Brief](#)

- Bauer HU, Schöllhorn W (1997). Self-organizing maps for the analysis of complex movement patterns. *Neural Processing Letters* 5, 193-199
- Button C., Davids K (1999). Interacting intrinsic dynamics and intentionality requires coordination profiling of movement systems. In: *Studies in Perception and Action V* (edited by MA Grealy and JA Thomson), pp.314-318. Mahwah NJ: Lawrence Erlbaum Associates
- Newell KM, Liu Y-T, Mayer-Kress G (2001). Time scales in motor learning and development. *Psychological Review* 108, 57-82

Published Dec 2001  
[editor](#)  
[©2001](#)