that directly exploit the possibility of hierarchical structuring of the index system.

### RESEARCH DEBATE

**Researchers, Abandon Statistical Significance Testing!**

Dr Ramakrishnan Menon  
School of Science, NIE

A renewed call

The call to abandon statistical significance testing (SST) and replace it with alternative statistical procedures has been renewed recently by Schmidt (19%). What makes this renewed call different from previous ones (e.g. Carver, 1978; Coats, 1970; Falk, 1986; Meehl, 1978; and Menon, 1993) is that:

- It is given by the 1994 president of the Division of Evaluation, Measurement and Statistics of the influential American Psychological Association.
- Some specific alternatives are discussed.

Examples are given of how reliance on SST retards the growth of cumulative research knowledge.

The alternatives suggested are (a) point estimates of effect size and confidence intervals for individual studies and (b) meta-analyses for multiple studies. The interested reader is strongly recommended to read Schmidt (1996) for further clarification.

According to Schmidt (1996), research based solely on SST usually goes through the following sequence:

- Initial optimism that social science research can answer important questions in the face of uncertainty and allow policy makers to implement certain changes.

The several studies conducted result in conflicting or non-conclusive results. SST research is undertaken to study the moderator or intervening variables, resulting in more conflicting results.

Research sponsors and the public become disillusioned and cynical of research enterprises.

- Researchers themselves begin to become cynical of their work.

Schmidt (1966) shows that these alternatives to SST will allow quantitative research methods to point to less contradictory research results and more constructive suggestions.

**Implications for NIECER and educational research in Singapore**

While Singapore prides itself on being competitive, on being an intelligent island, on spending millions on R & D, and on upgrading and developing professionally, researchers (both would-be and current) are still using and/or being taught SST, even though SST has been shown to be illogical and mathematically flawed (e.g. Menon, 1993). Recently, Schmidt (personal communication) stated that the American Psychological Association "now has a Task Force studying whether researchers should be discouraged from using significance testing in their research and in APA publications." Rather than waiting for the APA to take the lead in abandoning
the use of SST, could not NIECER and other bodies/associations in Singapore that have a research arm and/or journal (e.g. The Asia Pacific Journal of Education, The Mathematics Educator) be more proactive, by (a) actively discouraging the use of SST in research and research methodology courses and (b) rejecting articles that rely solely on SST for the credibility of their research results?

It is hard not to agree with Menon’s (1993) criticism directed at the misuse of statistical significance testing (SST) on logico-mathematical grounds.

Although Menon raised the issue in the context of mathematical education research, it has a long history of four decades in education research in general. The main harmful effects listed by Menon include the over-reliance on dichotomous SST decisions (leading to less attention being paid to the research design) and uncritical application of SST (leading to ritual rejection by journal editors of articles without ‘significant’ results, which in turn, leads to self-censorship by researchers). These are, in fact, the results of misuses of SST due to poor teaching and poor editorial policy more than the inherent weaknesses of SST as a statistical decision-making tool.

It is also hard to agree, on educo-pedagogical grounds, with Menon’s suggestion to cease teaching SST.

It does not take much creativity to imagine what may happen if research students are not taught SST. Since the literature review is an important, initial step in planning any study, research students will come across a plethora of studies reporting SST results. Without being taught SST, two lines of action are possible:

- To ignore SST results and wonder what they mean and signify. In the spirit of scholarship, this is to be discouraged.
- The conscientious research students learn about SST on their own, hopefully correctly. If the need to learn SST is real, why stop teaching it?

Over the years, several alternatives to SST have been suggested by its critics although there is still no agreement as to their usefulness. Menon, however, favours replication as the ‘most believable approach to a scientific method’.

There are three levels of replication: The direct, the systematic and the conceptual. It is not clear which of these Menon is referring to; perhaps all of them. Direct replication may be attainable in physical sciences where experimental control can be largely repeated exactly or almost so, but not likely nor desirable in educational research. Education is a multivariate situation where a very large number of

**References**


**Should we Execute the Victim SST?**

- A response to Menon

Dr Soh Kay Chang
Nanyang Academy of Fine Arts