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?

In doing so, not only will we be able to augment our research knowledge and help suggest some innovations in education arising directly from such research results, but we will also be one of the first to discontinue an outdated, illogical, flawed and slavishly-followed method of statistical decision making.

* * * * * *

References


Should we Execute the Victim SST? - A response to Menon -

Dr Soh Kay Chang
Nanyang Academy of Fine Arts

It is hard not to agree with Menon's (1993) criticism directed at the misuses 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 pedagogical grounds, with Menon's suggestion to cease teaching SST.

It does not take much creativity to imagine what may happen if SST 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
variables need to be studied and no one research study is expected to have the last say. Hence, as rightly pointed out by critics of SST, a study provides only a data point and to establish a theory, many such points are needed. Even if exact, direct replication is possible, information thus gathered may not be so useful as to warrant the extra effort.

This leads naturally to systematic replication where the original design can be kept more or less intact, so that earlier findings can be verified, and new variables included so that effects under different conditions can be investigated. This form of replication may be more consistent with the nature of educational research.

However, as educational measurement is not physical measurement, the international interchangeability of education and psychological measures is contentious at best. Researchers more often than not find themselves having to adapt measures developed by earlier researchers for their specific research environments. This suggests that conceptual replication may be more convincing and convenient when studying the same educational phenomenon across nations.

Systematic and conceptual replication will provide a wealth of information for theory-building when a sufficient number of related studies have been conducted. This, of course, can be attained through meta-analysis which not only accumulates knowledge but also explicates conditions in which the interested effect can be obtained.

It is a paradox that a probabilistic device such as SST, used for deterministic purpose, renders it an object of criticism while the real problem lies in its misinterpretation and misuse. Now that SST, as a victim, has been prosecuted not once but many times over, is it wise now to pass a death sentence and execute it? I should think not. On the contrary, I believe we should continue to teach SST and teach it better than we have been able to hitherto.

---

**DISSERTATION ABSTRACT**

**Relationship Between Peak VO\(_2\) and Performance on The Rockport Fitness Walking Test of Adolescent Males with Intellectual Disability**

Dr Teo-Koh Sock Miang
School of Physical Education, NIE

The 1-mile Rockport Fitness Walking Test, RFWT, has been recommended as a valid and reliable field test of cardiovascular fitness for adults with and without intellectual disability. However, it has yet to be determined if the RFWT is also a valid and reliable cardiovascular field test for children and adolescents with intellectual disability.

The purpose of this study was to determine the relationship between peak oxygen consumption and RFWT performance of 12-17 year old males with intellectual disability. Reliability of RFWT performance and peak VO\(_2\) was also evaluated.

Forty subjects (mean age = 14.13 years; mean IQ = 50) were selected from two special schools in Singapore. A three-phase familiarization and practice process preceded testing. All subjects were tested twice on the RFWT. Of these 40 subjects, 24 were tested twice on the graded maximal treadmill test (GMTT). Except for mean IQ, independent t-tests indicated no significant differences in age, weight and height of those who did the "RFWT only" and those who did the "RFWT & GMTT".

Test-retest reliability of the RFWT was high for walk times, RFWT peak End-Heart rate and peak RFWT HR (r = 0.90 to 0.97). Test-retest reliability of relative peak VO\(_2\) was 0.90 (mean = 41.28 ml · kg \(^1\) · min\(^{-1}\) ± 6.43). Relative peak VO\(_2\) correlated negatively and significantly with best walk time (r = -0.76). Correlation between absolute peak VO\(_2\) and walk time was -0.58. Partial correlation analysis indicated that when weight and various combinations of variables with weight were held constant, the relationship between walk time and peak VO\(_2\) was strengthened.

Multiple regression analysis of RFWT performance variables and peak VO\(_2\) measures indicated...