A New Latent Trait Modeling Approach with Contributions to Assessing Latent Unidimensionality and Ability Test Bias

The objectives and progress of W. Stout on the 1984 ONR contract are summarized with references to reports, unpublished works, and published papers detailing this work. Important accomplishments include: (a) correction of Stout's statistical test of unidimensionality for presymptomatic performance bias, (b) a large scale simulation study suggesting that the statistical test should perform well in applications, (c) development of a new IRT modeling approach capable of assessing the number of dominant latent dimensions and, (d) a rigorous definition of ability test bias and a statistical test of ability test bias.
The purpose of this final report is to summarize the work accomplished by Stout and his graduate students as a result of the 1984 Office of Naval Research grant. The work accomplished by Humphreys and Tucker and their graduate students as a result of the 1984 Office of Naval Research grant is described in a companion final technical report.

Major Research Objectives of Stout:

(a) Convert Stout's statistical test of psychological test unidimensionality into a useful procedure having good statistical properties.

(b) Do a large-scale realistic simulation study of Stout's statistical test of unidimensionality in order to assess its performance in test settings typical of actual psychological ability tests.

(c) Develop a sound theoretical model for psychological test bias and a good statistical test of psychological test bias.

(d) Develop a better Item Response Theory modeling approach than that afforded by assuming local independence: in particular, be able to model the existence of one or more dominant dimensions in the presence of possibly numerous minor dimensions.
Progress: Prior to the inception of the 1984 project, Stout had developed a statistical test for the unidimensionality of an ability test. This statistical procedure was defended by an asymptotic statistical theory. This work is summarized in the June 1984 Office of Naval Research technical report:


However, a pilot simulation study had shown that unless the test was very long (≥ 80 items), the performance of the procedure breaks down and becomes biased in the sense that unidimensional test data is judged much too frequently to be multidimensional. One of the major accomplishments of the 1984 project was the development of a sophisticated bias correction for the Stout statistic. This correction produced a hypothesis testing procedure that accurately judges unidimensional test data to be unidimensional even for short tests (as few as 25-30 items) in the presence of large numbers of examinees with guessing assumed.

A very large scale "realistic" simulation study was conducted to better assess the performance of the bias corrected statistical test. Indeed, nearly 6300 test administrations were simulated. The study showed conclusively over a wide range of realistic settings that the statistical procedure adheres well to the nominal level of significance given by the asymptotic theory while displaying excellent power against multidimensional alternatives. A thorough report of this simulation study and of the bias corrected statistic and the asymptotic theory for it is presented in:

R. Nandakumar, former Ph.D student of Stout now at the University of Delaware, discovered that the statistical test of unidimensionality sometimes performs poorly in the presence of highly discriminating easy items with guessing assumed. Nandakumar then developed a modification of Stout's statistic and showed through a simulation study that the modification nicely corrects the problem. Nandakumar further modified the manner in which Stout's procedure selects "assessment test" items, allowing for a random number of such and hence improving the power of the statistical test. This research, plus other work, is described in


The foundations of a new and hopefully more realistic Item Response Theory approach were a major development of the 1984 project. In this, the traditional notion of unidimensionality is replaced by a new notion of unidimensionality called essential unidimensionality. Essential unidimensionality counts only the number of dominant latent dimensions while ignoring the number of minor dimensions. As co-investigators Humphreys and Tucker have emphasized in their writings, the existence of possibly numerous minor dimensions is an inescapable empirical fact of ability testing. It is shown that Stout's statistical test of unidimensionality described above is formulated to precisely assess essential unidimensionality. In this sense, Stout's test of unidimensionality is model based. Much of the new theory, including some refinements made after the expiration of the 1984 grant, appear in the Office of Naval Research technical report:

A modified version of (iv) has been accepted for publication in Psychometrika.

Growing out of the above-discussed new Item Response Theory modeling approach, Stout has proposed a rigorous definition of ability test bias that explicitly takes into account that biased tests are multidimensional objects designed to measure a particular dimension -- for example, mathematics ability -- but with one or more other dominant dimensions present -- such as would be created by a large proportion of mathematics items having a strong verbal component. This is described in the technical report (iv) mentioned above.

R. Shealy, a Ph.D student of Stout's who will receive his Ph.D in Fall 1988, has used the above definition of test bias to develop a statistical test of test bias. This procedure is defended by an asymptotic statistical theory and a large-scale simulation study. It will be reported on in a future technical report.
END DATE
10-88
DTIC