

The anglicization of American personality tests: Rejoinders

Scientific rigour vs scholarship: A Response to Cook and Gough

Reading the response prepared by Cook and Gough, to my article on the anglicization of American personality tests, reminded me of the character Humphrey, in my favourite TV series, *Yes, Prime Minister*. Humphrey, supposedly the Prime Minister's civil servant, specializes in obscuring the truth behind simple issues by using unnecessarily complex language. Sadly, in this TV series, obfuscation often wins the day, passing off as erudition and scholarship. However, as most students will know, true scholarship belongs to those (pedagogues) who explain complex issues in ways that makes these issues appear deceptively simple. Therefore, for the benefit of Cook and Gough, I will briefly restate the main points of my article.

1. Until very recently British test publishers have been routinely selling American personality tests for use in the UK, with little apparent concern about problems of 'cross-cultural validity'.
2. These same test publishers have recently done a U-turn on this issue, and have suddenly decided to publicize concerns about the problems of using tests that may lack 'cross-cultural validity'. (While there may be good commercial reasons for this sudden volte-face, the objective, scientific reasons for this are unclear.)
3. Rather than set objective, verifiable criteria for assessing the cross-cultural validity of

Richard Budd

personality tests publishers have, instead, chosen to 'anglicize' American tests in an ad hoc manner. (From a scientific perspective, this is not scholarship!)

4. The ad hoc manner in which test publishers have 'tinkered' with these standardized psychometric instruments means that test users can no longer be sure that the 'anglicized' forms of these tests are measuring the same constructs as the original (American) forms of these tests.
5. Test publishers have provided no data demonstrating that the anglicized forms of these tests are equivalent to the original (American) forms of these tests. However, they have continued to cite validity studies (collected on American samples, using the original forms of these tests) as evidence of the validity of the 'anglicized' forms of these tests.

I will leave the reader to decide whether these points have been addressed by Cook and Gough, and whether their appeals to scholarship are no more than an attempt to obscure their lack of scientific rigour.

A healthy debate continues: A reply to Lord and Smith

I greatly enjoyed reading the response by Wendy Lord and Pauline Smith, who raise a number of

very important points, with which I wholeheartedly agree. As they rightly note, if multinational companies are to use the same personality tests across all their sites, then it is important that these tests demonstrate high levels of cross-cultural validity. Moreover, even when such tests are used within national boundaries it is still important that they are cross-culturally valid, as test respondents often come from a variety of different cultural backgrounds. Thus it is important that test publishers routinely provide data (within test manuals) that demonstrate the cross-cultural validity of the instruments they sell. As Lord and Smith point out, adequately addressing problems of culture-fairness is not only a scientific issue, but is also an ethical responsibility for all test users.

To this end, there is a need to elaborate agreed methods for developing tests that are valid across different cultures, and to develop procedures for assessing tests' cross-cultural validity. As psychometrics is an area of scientific research, such methods need to be both objective and verifiable. This goal is only achievable if such methods are grounded in theory; be it classical test theory, or item response theory. Clearly, little is to be gained by approaching this important issue in

an idiosyncratic, subjective manner; even if the intention behind such *ad hoc* attempts at anglicization is laudable.

As I have already argued, making such *ad hoc* changes to standardized personality tests can inadvertently generate a variety of potentially serious problems with these tests' sub-scales. Even when attempts at item anglicization are undertaken by 'expert test developers', surprises can occur. For example, I recently received a data set for the original (American) version of the 16PF-5; completed by a large sample of British respondents. Examination of Factor I indicated an average corrected item-whole correlation of 0.37; indicating a reasonable degree of item 'cohesiveness' for this scale. To my surprise I discovered that the previously cited item concerning a preference for dancing or fencing, to wrestling or baseball, as forms of exercise, had a corrected item-whole correlation of 0.37. This indicates that in spite of reservations about whether British subjects understand this item, it appears to work just as well as do items with no apparent 'Americanisms'.

Thus, like many others, I fell into the trap of trusting my 'expert judgement' rather than relying upon objective data - *mea culpa!*

Help!

Most material submitted to *Selection and Development Review* has been prepared on a word-processor of one form or another, but often we only receive a printed copy. This means that the text either has to be retyped or scanned in electronically - both methods taking up time and increasing the risk of mistakes creeping in. So, when sending in items, **please, please, please:**

- ☐ Provide a clear, up-to-date printed copy.
- ☐ Enclose a 3.5in disk (either DOS or Mac format) with the document saved both in its original wordprocessing format and as an ASCII file.
- ☐ Save diagrams and other illustrations both in their original format and as a TIFF or EPS.

This will only take you a few seconds but saves literally hours when SDR is being laid out - and it reduces the risk of errors.

Confused? Ring Geoff Ellis or Mark Bassett at the Society office (0116 252 9523) for assistance.

Many thanks