Professor Barbara Griffin
Professor
Department of Psychology
Macquarie University, Australia

Selecting Medical Students: What are We Measuring and are We Measuring it Consistently?

Abstract:

It has long been accepted that successful performance as a medical student or medical practitioner requires both cognitive and non-cognitive skills, with identified core skills including academic ability, integrity, the ability to relate interpersonally, and good decision-making. It is therefore increasingly common for students to be selected not only with the use of a cognitive ability test, such as the UMAT, UKCAT, or MCAT, but also with interviews, personality tests or situational judgment tests designed to assess non cognitive aspects.

Justification for including these different tests must be provided by evidence of predictive validity. I therefore begin by presenting new data on how cognitive tests and multiple mini interviews (MMIs) contribute independently to student performance *across time*, distinguishing between those students who maintain average or above average academic performance over a medical degree with those whose performance is consistently below average.

Despite this, and a larger body of research showing reasonably good predictive validity, at least for MMIs and academic or cognitive ability, construct validity remains somewhat unclear. In other words, we are still uncertain as to exactly what such tests are measuring. Studies of the UMAT test and selection interviews are used in this presentation to illustrate the complexity of identifying the constructs being assessed by selection tests. In particular, by comparing scores of applicants who were interviewed at multiple institutions, it becomes apparent that interviews are less likely to assess a range of 'narrow' traits, such as empathy, communication and team work, and are more likely to assess one overall latent construct. Possibilities for this underlying factor are 'potential to be a good doctor' or 'professionalism,' but MMIs and panel interviews may actually assess a different construct.

Of concern however, is that in the high stakes context of medical student selection there are external influences that threaten construct validity. These include repeat testing and commercial coaching. Coaching appears to reduce the extent that tests like the UMAT measure cognitive ability by increasing 'test-wiseness', which in turn limits predictive validity. Repeat testing appears to improve both cognitive ability and interview scores but such strong practice effects are also like to alter the constructs being assessed. Whether or not improved scores increase one's chance of selection may depend on how they are used to rank applicants.