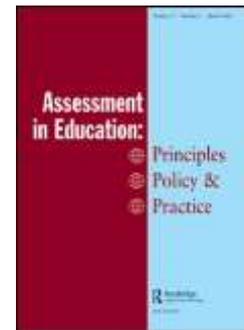


Assessment in Education

High-Stakes Testing – Value, Fairness, and Consequences

Guest Editors: Gordon Stobart and Theo Eggen

An Audio Clip with Gordon Stobart, an Executive Editor of *Assessment in Education* and Special Issue Co-Editor



Hello, I'd like to introduce a Special Issue of *Assessment in Education* on the impact of high-stakes testing. I'm Gordon Stobart, Emeritus Professor of Education at the Institute of Education, London, and my Co-Editor is Professor Theo Eggen from the University of Twente in the Netherlands. We chose high-stakes testing and the impact of high-stakes testing because it dominates so much educational policy and practice across the world and we wanted to examine this impact both positive and negative.

What do we mean by high-stakes testing? We take a test to be high-stakes when it has serious consequences for at least one stakeholder. It could be the student, it could be the school, the local authority, the government, or it could be all of these. Historically speaking high-stakes testing has primarily been about selection, so it's been high-stakes for the candidate because of the life chances that come from the results. This use goes way back 2,000 years to the Chinese Civil Service Exam, and the appeal from then onwards has always been its fairness in comparison to other means of selecting people which were usually to do with patronage, family connections, and the like.

Several of the papers in this Special Issue consider fairness, particularly when there's pressure to get unfair advantage through cheating, manipulating the system, and the like. So, for example, we've got a fascinating paper on how high-stakes tests for university selection have developed in the former Soviet Republics, where selection was largely based on political patronage previously. Here the emphasis is on keeping the exam process secure and reliable, and the extent to which examination bodies have to go to do this.

We've also an article from the Caribbean that looks at 11+ selection, which historically has been one of the most life changing selection exams, very high-stake, and they look at another aspect of fairness, the idea of gender and social bias, and what happens when an exam is modified in order to try and account for some of this and some of the unintended consequences.

Unintended consequence is the theme of a study of dramatic change to a high-stakes selection exam for university entrants in Cyprus. A case study of what happens when a long-established exam system is suddenly replaced, this was because of European Union requirements, by a different one. These papers cover the use of high-stakes testing primarily for selection purposes.

A major development over the last few decades has been the use of testing for schools, for accountability for schools, local authorities and states being accountable to government. This has often taken the form of external tests which provide data on achievement at school, local/state level. Many will be familiar with the debates around the impact of national tests in England and in the United States with the 'no child left behind' legislation, where schools are judged by the results,

with serious consequences if the results do not meet the targets that have been set for them. So the issue of teaching to the test and the way schools prepare students for the test becomes important.

Despite what we know about the impact, often negative, of such tests, there are still countries that are continuing to develop high-stakes tests for national accountability and we have a case study of such – the recent introduction of national testing in Australia. An even more recent form of accountability testing has come through international comparisons such as PISA, *Programme for International Student Assessment*. The significant factor about these is that they are a sample based assessment and therefore low stakes for individuals and schools, no consequences for them, but become high-stakes for politicians and policy makers. If a country does well or badly in such as PISA there are all sorts of policy consequences and we include in this issue a powerful example of this from South Africa, which bravely took part in the TIMS study, the *Trends in International Mathematics and Science Study*. What happens when the results are very disappointing, and we see the serious consequences here for policy makers and the education department in South Africa.

These studies all show the importance and consequences of high-stakes testing. Given this, such tests need to be high quality ones. We include two papers that look at the issue of quality in the design and use of such tests. The first from Canada looks at how we validate such tests and offers a rethink of what validity involves when the results, in this case state-wide maths tests, are used for several purposes. The second technical paper looks at the quality of decision making in the important Secondary School Leaving Diploma in the Netherlands. This Diploma involves aggregating a wide range of assessments into a single pass/fail decision. The paper examines how this can be done in the fairest possible way.

We found putting together this Special Issue both stimulating and informative given the richness of these international contributions. We hope you find the same too. Thank you.