

0728

Vagueness and Sincerity in High Stakes Educational Assessment.

Richard Marshall

Institute of Education, London, United Kingdom

Background

Part of what makes an assessment competent is that the assessment is sincere. Sincerity is an epistemic condition placed on the assessment: it is about believing in the assessment made, about having confidence in it. The integrity of the system depends in part upon this. An assessment system that sorts candidates into degrees of success and failure is prototypical. Such systems have borderlines between grades that are sharp. However, a reason for thinking that the sharpness of the borderline is unknowable is that grading systems are impossible objects. If this is so then any claim to know the borderline has to be insincere.

Research Questions

The paper looks at these claims and concludes that any assessment system that claims to know where its sharp borderlines are between grades is an insincere system. When assessments are for high stakes sincerity is a key issue. The paper considers the predicament of an assessor having to make a decision in a borderline case for high stakes. There is evidence that success in summative exams in the UK increases average earnings and life expectancy rates. High stakes tests change what is taught and how it is taught.

Methods

Much literature discusses the impact of 'high-stakes' tests and asserts that they lead to four negative things : teachers focussing on the content of the tests; the frequent administration of practice tests; the training of students in answering test questions to the exclusion of genuine learning; the adoption of 'transmission' teaching style, each of which prevents genuine learning. On top of this, the role of school as a transmitter of values and practices was also becoming less effective than required as more pupils in the system found schooling increasingly alienating. 'Throughout the 1990's, evidence was accumulating of the detrimental effect of frequent testing on students' enjoyment of school, their willingness to learn, other than for the purposes of passing tests or examinations and their understanding of the process of learning.' (Harlen 2005)

Frame

Much of the literature relates to issues of reliability and validity of assessments and covers the different approaches that have been taken to achieve these. Other literature relates to the link between the high stakes of the assessments and the impact this has on the kinds of assessments used for this purpose.

Having looked at the issues in this literature the paper then introduces philosophical vagueness into focus in order to raise the issue of sincerity to the discussion. Vagueness is largely an area that has not explicitly appeared in the educational literature before, with a few honourable exceptions, but is a well established area in philosophical literature. The paper also looks at the way the issue has been treated in law, another profession where making high stakes judgments is a core business.

The paper explains the reasons for thinking that high stakes assessment systems are impossible objects by looking at the literature about philosophical vagueness. Grading systems are compared to colour spectra. This paper is suggesting that we current grading systems for high stakes tests developed and refined over the last fifty or so years, as described above, are impossible objects like a colour spectrum of red to non-red. We can imagine why the paper believes this if we simplify the

situation of grading. Imagine a system where what is being tested is the smartness of candidates. 'Smartness' is understood in a non-technical sense that everyone understands, both in and out of the education system. The system has invented an assessment that can validly and reliably detect smartness. It recognizes that there are degrees of smartness and that candidates range from very smart to not smart at all. Just like a colour spectrum can capture degrees of red running from very red to not red at all, our grade system of smartness is equally capable of showing degrees of smartness running from smart to not smart.

Of course, the big difference is the source of the beliefs. A colour spectrum reflects perceptual belief. A smartness spectrum reflects beliefs about smartness that are derived from knowing what smartness entails. Controversy of course surrounds the nature of smartness, but for the sake of the illustration, let us assume that there is general agreement about smartness and how we know it. However this is accomplished, assume everyone agrees that the grade spectrum captures accurately the knowledge of smartness and its degrees.

Research findings

The difference between the colour and the smartness spectra is not that they don't both have a transition between red and non-red, smart and non-smart. They do. The difference is that the grade spectrum claims knowledge of where the transition between smart and non-smart is. Where the colour spectrum seems fuzzy at the borderline, so that it isn't clear where the border is, or whether there even is a border, the grade boundary used in the schools is precise and sharp.

The strangeness of this is often overlooked. The strangeness lies at the heart of this thesis. We can imagine a teacher presented with two candidates work that are different by only a single mark. The teacher is unable to discriminate between the two candidates because the difference is too small to be noticed. Yet the grading system, which is supposed to truly reflect 'smartness', is able to make the distinction. The strangeness derives from this fact. Imagine a situation where everyone in the world is asked to differentiate between the two candidates and no one can, then this merely emphasizes the problem. A distinction is being made that no one on earth can recognize or make.

The assessment system able to make distinctions that no human can make can be likened to an inhuman or superhuman machine, transcending the limited powers of fallible humankind to deliver truths our minds cannot grasp.

Vagueness that creates these absolute borderline cases means that all decisive judgments in these cases are insincere. Judges, teachers and assessors who make decisions in these cases are lying.