Thinking Skills and Creativity 10 (2013) 265–279

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: http://www.elsevier.com/locate/tsc

The 21st century challenge for science education: Assessing scientific reasoning

Jonathan Osborne *

Graduate School of Education, Stanford University, Stanford, CA 94305, United States

ARTICLE INFO

Article history:
Received 12 February 2013
Received in revised form 3 July 2013
Accepted 19 July 2013
Available online 7 August 2013

Keywords:
Scientific reasoning
Assessment
Higher order thinking
School science

ABSTRACT

Expectations of the outcomes of education in the 21st century increasingly focus on higher order thinking of synthesis, analysis and evaluation. Yet school science education is still dominated by lower level cognitive demands—in particular recall. The argument made by this paper is that the failure to transform science education for the needs of the 21st century is a consequence of a lack of a good model of scientific reasoning and a body of expertise about how to assess such higher order cognitive competencies.

In response, this paper presents a model for scientific reasoning which is a synthesis of contemporary philosophical perspectives and empirical psychological studies of how scientists work. Such a model offers some insights into the kind of competencies that science education might seek to develop to address the contemporary demands of society. Scientific reasoning is, however, domain specific and dependent on a knowledge of the content and concepts of science; a body of procedural knowledge about standard methods; and an epistemic knowledge of how such procedures warrant the claims that scientists advance. Assessing 'what counts' depends on a deeper understanding of what counts—in this case the nature of the performance and the knowledge base required for the display of higher-order thinking reasoning.

Finally, it is argued that recent developments in computer-based platforms such as the open-source TAO platform to be used for the PISA assessment in 2015 and other computer-based platforms offer the promise of enabling students to display a wider range of performances and more sophisticated methods of assessments. Better assessments are not possible, however, without better constructs and likewise, better constructs cannot be assessed without a broader repertoire of methods of assessing student performance.

1. The common rationale for science education

Commonly, teaching science in school has been justified by an economic rationale. For instance, the rapid and radical transformation of society in the first half of the 20th century—in particular its contribution to victory in two world wars strengthened initial arguments for the importance and value of science education (Committee to Enquire into the Position of Natural Science in the Educational System of Great Britain, 1918; Layton, 1973). Since then an economic rationale for science education has been an enduring feature of the debate about its value and purpose and repeated regularly forming the basis of the Dainton report (Dainton, 1968) in the UK in the 1960s; arguments that the US nation was at risk of losing the

* Tel.: +1 650 725 1247.
E-mail address: osbornej@stanford.edu

1871-1871/$ – see front matter
http://dx.doi.org/10.1016/j.tsc.2013.07.006

economic competition with other nations in the US (National Commission for Excellence on Education, 1983) in the 1980s; and then repeated in the last decade both within the US (National Academy of Sciences, 2010; National Academy of Sciences: Committee on Science Engineering and Public Policy, 2005), the UK (Lord Sainsbury of Turville, 2007; Roberts, 2002) and Australia (Tytler, 2007). In all of these reports the pressing issue of the economic imperative is bluntly presented and posed as a threat to the nation's well-being as a rationale for improving science education and the system that supports the STEM infrastructure.

As a result, over the years school science has acquired the mantle of a pre-professional training for the future scientist such that their needs have come to dominate the curriculum. Indeed, educating the future scientist is best served by an education which is foundational and builds a knowledge of the basic concepts and language of the discipline (Millar & Osborne, 1998). In such an education, an understanding of the overarching conceptual coherence and the nature of the discipline itself only emerges for those who complete undergraduate, if not graduate education (Bøe, Henriksen, Lyons, & Schreiner, 2011; Lyons, 2006). To the novice lacking any overview science can too often appear to be a 'miscellany of facts' (Cohen, 1952) akin to being on a train with blacked out windows where only the train driver knows where you are going (Claxton, 1991). Absent from this form of education is any development of student facility with critique (Ford, 2008; Ford and Michael, 2012)—the essential activity which requires the higher order reasoning of comparison, contrast and evaluation. Schwab famously argued that one consequence was that science was taught as a 'rhetoric of unmitigated conclusions' (Schwab, 1962); Duschl (1990) too argued against the common practice in science education of presenting 'final form' knowledge as unequivocal and uncontested—a practice which lead Horton (1967) to argue that

> . . . the grounds for accepting the models proposed by the scientist is often no different from the young African villager's grounds for accepting the models propounded by one of his elders. In both cases the propounders are deferred to as the accredited agents of tradition. p. 209

A second consequence is that students emerge from their education believing that knowledge that has the status of a fact is the apotheosis of scientific achievement (Driver, Leach, Millar, & Scott, 1996) when, in contrast, it is 'theories which are the crown of science' (p. 168) (Harré, 1984). For instance, overwhelmingly the names of scientists that exist in perpetuity – Darwin, Einstein, Wegner, Maxwell, Copernicus, Newton etc – are all recognized for their theoretical contribution and their ideas which have profoundly changed our conceptions of the material world. A third consequence is a pedagogy dominated by transmission with an absence of arguments from evidence for scientific ideas (Driver, Newton, & Osborne, 2000; Kuhn, 2010) where science is presented as 'as static bodies of knowledge, focusing on vocabulary and algorithms' (p. 46) (Weiss, Pasley, Sean Smith, Banilower, & Heck, 2003). In their extensive observational survey of the teaching of science and mathematics in US, Weiss et al. (2003) found that only '14 percent of lessons nationally having a climate of intellectual rigor, including constructive criticism and the challenging of ideas' (p. 54). Hence, Rogers' comment that 'we should not assume that mere contact with science, which is so critical, will make the students think critically' (Rogers, 1948) is still as true today as it was then.

The thesis of this article is that the practice of science education has failed to come to terms with the vision offered by Hill (2008) that the societies that sustain their competitive edge in the coming decades will be 'post-scientific' societies. In such a society, highly valued competencies will be the ability to draw on a range of disciplinary knowledge, and notably, to think creatively and evaluate new ideas in a critical, reflective and rational manner. Hill argues that employers will require individuals who, while having a core understanding of scientific and technical principles, have the ability to communicate and synthesize knowledge in an original manner. Likewise, experts invited to an NRC symposium in 2007 to explore the nature of the skills required for the future workplace argued that the competencies that would be demanded were, among others, the ability to solve problems creatively, sophisticated communication, self-management and systems thinking (National Research Council, 2008). Gilbert puts it even more straightforwardly arguing that 'in a world where there is an oversupply of information, the ability to make sense of information is now the scarce resource' (Gilbert, 2005). More recently, the NRC in their report on Education for Life and Work (National Research Council, 2012a) argued that it is important to develop three domains of competence—the cognitive, the intrapersonal, and the interpersonal. Central to the first of these is the development of students' ability to undertake the cognitive process of complex reasoning which includes critical thinking, non-routine problem solving, and constructing and evaluating evidence based arguments. Taken together, such arguments suggest that the valued competency and ability of the future will be higher order reasoning of evaluation, synthesis and critique. However, as Hill argues, if these are to be a feature of science education, then it is important to 'be certain that we emphasize what we want, for we shall surely get what we emphasize' (Hill, 2008).

Even the economic imperative is demanding more of contemporary education – particularly science and mathematics – than it currently delivers. In a series of papers drawing on the data available from international test initiated originally by the IEA, Hanushek, Peterson and Woessmann (2010); Hanushek and Woessmann (2011, 2012) show that, after controlling for years of schooling and initial GDP per capita, there is a statistically significant relationship between test scores and growth in GDP per capita between 1960 and 2000. Indeed, 'test scores that are larger by one standard deviation (measured at the student level across all OECD countries in PISA) are associated with an average annual growth rate in GDP per capita that is two percentage points higher over the whole 40-year period (p. 638).' Another way of interpreting this finding is that raising the achievement of all students by 25 points (1/4 of a standard deviation) on the PISA tests would lead to very significant gains in economic growth.