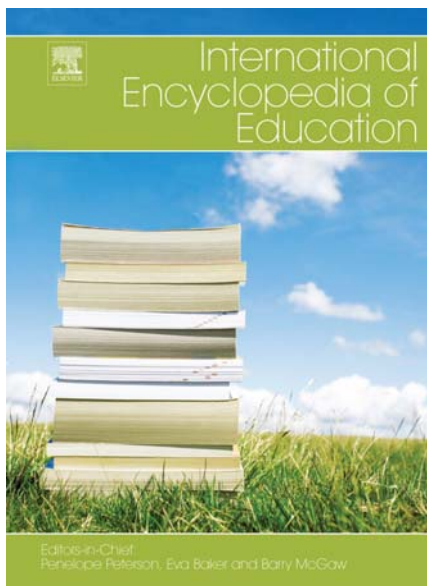


Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This article was originally published in the *International Encyclopedia of Education* published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

Gardner J and Holmes B (2010), Students' Rights in Assessment and Evaluation. In: Penelope Peterson, Eva Baker, Barry McGaw, (Editors), *International Encyclopedia of Education*. volume 3, pp. 378-383. Oxford: Elsevier.

Students' Rights in Assessment and Evaluation

J Gardner, Queen's University Belfast, Belfast, Northern Ireland

B Holmes, Hamdan Bin Mohammed e-University, Dubai, United Arab Emirates

© 2010 Elsevier Ltd. All rights reserved.

If rights are interpreted strictly as legally supported entitlements, enacted through statute, then this article should probably end at this point. As far as we are able to determine, there are no specific rights for students in relation to the assessment or evaluation of their learning in any nation's legislation. (Arguably, the terms assessment and evaluation are nuanced in meaning but for the most part, assessment will be used in the article to relate to any process designed to appraise student learning.) However, there do exist educational principles, widely endorsed standards and codes of practice, and at least one legally binding set of international articles that can be used to underpin a discussion of students' rights in assessment.

Most jurisdictions have education-related rights based on statutory or common law, which may apply to assessment contexts depending on the circumstances. Some of these rights have considerable case law, most notably in the US and to a lesser extent the UK, Europe, and Australasia. They cover issues such as discrimination (numerous instances of disability, race, and gender actions over the past half century), parents' versus children's rights, free speech on campus, and youth culture. However, the most amenable measures, which afford at least the potential for legal protection of students in contexts involving assessment, are espoused under two internationally respected codes: the Standards for Educational and Psychological Testing (AERA *et al.*, 1999) and the Code of Fair Testing Practices in Education (JCTP, 2004). A third axis of potential protection derives from two articles of the United Nations Convention on the Rights of the Child (UNCRC, 1990), namely articles 3 and 12. These standards, code, and articles are discussed below, but first it is worth outlining why students should have the right to a greater say in the assessment of their learning.

Purposes of Assessment

Assessment can be argued to have many purposes, roughly positioned along a spectrum that ranges from the ad hoc question-and-answer sessions of teachers and students in classrooms to the formal examination style of assessment that tests students' levels of achievement after a period of learning activity. The former is generally located at what is called the formative end of the spectrum, where the purpose is primarily to provide support for students' learning, while the latter is generally located at the

summative end of the spectrum, where the purpose is primarily to identify a measure of the students' learning.

Assessment as Support for Learning

Over the past 10–15 years, there has been a growing awareness that the primary role of assessment should be to support learning. This realization is emerging as the result of a number of key developments, most notably the rise in importance of formative assessment, or assessment for learning as it has become known. A growing body of research-informed theory and knowledge has been accumulating in the field of formative assessment in relation to its practice and implications for policy (Gardner, 2006). In common with pedagogical approaches such as dynamic assessment and dialogic teaching, assessment used as support for learning promotes a variety of processes involving increased student participation and their owning of the learning activities. These include: sharing in the process of identifying learning objectives and success criteria between the teacher and the students; the development of peer and self-assessment; effective questioning, listening, and feedback techniques; and a conscious determination on the part of the teacher to recognize and affirm the full range of student achievement. Such processes promote the students' voice in their learning and assessment, while seeking to maximize engagement, not just with the lesson activities and content but with each other, with the teacher, and with the learning process itself.

The theory of formative assessment eschews the traditional didactic model of passive learners and directive teachers, instead giving the students their say in important aspects of their learning and its assessment. The argument draws on widespread evidence that formative assessment approaches can contribute to higher levels of student motivation, achievement, and participation.

Assessment of Learning

At the other end of the assessment spectrum, assessment of learning, which is generally summative in design and intention, largely continues to illustrate the phenomenon of doing to rather than with students. Such summative assessments may have several purposes but generally speaking they come at the end of a period of learning

(end of semester, end of year, etc.), are formal compared to classroom assessments, and purport to be objective and trustworthy. The underlying traditions of assessments for summative purposes include isolating students from sources of advice, support, or knowledge, and requiring them to address a set of tasks (questions, essays, problems, etc.) in time-bound circumstances, usually by means of pen-and-paper tests. Such conditions do not readily enable students to participate in the design or the content of the assessment. However, once the work is assessed, the students do enjoy some widely accepted entitlements, including specified circumstances in which they can seek to have grades or scores reviewed. In most national-education systems today, the main purposes of summative assessments may be summarized under two headings: accountability and credentialing.

Accountability Purposes

There is a variety of purposes to which assessments may be put and the most topical are arguably those that serve accountability. Aside from the traditional accountability of learners (how well they have performed), the last couple of decades has witnessed vigorous programs in the US and UK that incorporate the assessment of student learning to hold teachers, schools, and school authorities to account. One of these, the hugely important No Child Left Behind Act (NCLB) 2001 in the US, is silent on students' rights or perspectives on any aspects of its processes or impact, including assessments, though parents' rights are exhaustively detailed and secured. It has been in operation for 6 years without any noticeable attention to the students' perspectives on its processes or impact. The US NCLB and the UK Making Good Progress programs are typical of many recent educational policy developments around the world, inasmuch as they set targets for schools, based on the dubious proposition that testing and increased accountability raises standards. NCLB has attracted from educationalists who see it as damaging to its main disadvantaged target groups and to individual students (e.g., [Houston, 2007](#); [Nichols et al., 2005](#)).

It is only a small step from local accountability to making whole nations accountable for the success of their education provisions through, for example, international comparative programs such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA). Almost without exception in these cases, the students are the cannon fodder in a battle between policy-makers and the school communities they oversee. They have next to no voice (except perhaps through the ventriloquism of their advocates) and precious little in the way of rights. [Houston \(2007\)](#), for example, claims that one of the deadly sins in the US NCLB is the "subjecting

of children to days of examinations annually, with the time taken away from instruction" while [Skidmore \(2003: 34/35\)](#) in the UK cites the statistic that "... an average pupil [in England] will now take 70 exams by the time they finish school [at 16]." In none of these circumstances has the student any rights to constructive participation in, for example, contributing to the identification of assessment criteria or design of assessment items and tasks. They are merely the subject of the tests.

Doing all of this to students has considerable cost implications. For example, the UK Association of College and School Leaders claims that many schools in the UK are spending more on examination fees than on learning resources, and that costs for examinations have risen by 51% in the past 3 years.

PriceWaterhouseCoopers ([PWC, 2004](#)) has modeled the examinations system in England (ranging across a wide variety of testing programs) and reported that in the year 2003–2004, its direct cost was £370 million with an additional £240 million on time costs, but excluding teachers' time in their assessment-support activities. There were also huge environmental costs in, for example, the distribution to test centers of 4.2 million test papers, 80 000 packages, and 237 million A4 sheets ([Durant, 2003](#)). In the US, [Wallis and Steptoe \(2007\)](#) estimate that the costs of independent state testing programs in the US (separate to National Assessment of Educational Progress (NAEP) testing) is in the region of US\$600 million.

Although the costs of external assessments may give cause for concern, the justification for such a huge investment in testing appears to be its perceived reliability. What, then, if external tests are not as reliable as the confidence in them implies? Research shows that systemic errors arising from the difference between students' observed score on specific test-taking occasions and their true score, as would be determined by an infinite number of test-taking occasions, can have considerable impact on results. These standard errors of measurement can cause an astonishing degree of misclassification in high-stakes testing ([Black and Wiliam, 2006](#)). In 2006, some 65 000 UK General Certificate in Secondary Education (GCSE) results were appealed and over 14 305 were regraded. Though a small percentage of the huge number of individual examinations taken (6.5 million), the fact is that many more students were likely to have been misgraded but they either did not know this or were unaware or did not take up their right to a review.

The most important observation in this situation is that these students exercised their right to request a review of their grading. This right has been offered by the UK examination authorities to students for some time. However, to assume that such a right is granted primarily from a natural-justice perspective would probably be an overly generous analysis. [Newton and Whetton's \(2002\)](#) four objectives for student appeals illustrate that the right of

appeal has less of a rights perspective than a pragmatic purpose. They identify appeals as serving a measurement objective, which is designed to maximize the measurement accuracy of the test in question (i.e., more accurate grade attribution) and a political objective, which is intended to safeguard public and professional confidence (i.e., the “if we find a mistake, we fix it!” assurance). There is also an educational objective for student appeals, one of contributing to teachers’ professional development by sharing the outcomes and reasons for appeal decisions. Finally, there is a tip of the hat in the direction of students’ interests with the objective of aiming to minimize any negative psychological consequences for them.

Credentialing Purposes

On an individual level, the credentialing purpose of assessment, that is, the award of qualifications in academic or vocational settings, is perhaps the most pressing for many students. Test programs that lead to qualifications or credentials are generally taken by students because they need them to demonstrate that they have a certain set of skills or that they have the necessary qualifications to progress to the next step in their careers. For some years now, qualifications-oriented programs have been expanded to include key skills (sometimes known as basic or transferable skills) such as literacy, numeracy, information technology, and working together. The concept of twenty-first-century skills is also gaining ground, for example, Baker’s “... adaptive problem-solving, risk assessment, managing distraction, self-management and changeable roles” (Baker, 2007). Such skills are simply not being assessed properly at present, if at all, and much needs to be progressed before students can avail of their entitlement to valid and reliable assessment in important new areas of learning.

The pursuit of qualifications may also have the objective of seeking admission to a college or university course, to a selective school or to the world of employment. As such they are high-stakes tests or assessments and it is entirely reasonable for students to expect and indeed demand accurate and valid assessments of their performance. To this end, two codes of good practice, the Code of Fair Testing Practices in Education (JCTP, 2004) and the Standards for Educational and Psychological Testing (AERA *et al.*, 1999), have been widely endorsed as guides to good assessment practice.

Writ large in the American Educational Research Association (AERA) standards is a recognition of the perspectives of the different stakeholders in assessment contexts; from parents, teachers, and students through to the test developers and those who use them for assessment purposes. Most importantly, for our purposes, they also provide a comprehensive overview of the rights and

responsibilities of students (or test-takers as the standards more precisely define those who are assessed) whether in educational, legal, or employment contexts. For example, the first reliability standard, Standard 2.1 (S 2.1) states that:

... estimates of relevant reliabilities and standard errors of measurement or test information functions *should* be reported (our emphasis) (AERA *et al.*, 1999: 31).

It would seem reasonable that students should be able to determine the accuracy of any measurements of their performance, but in some significant instances, most notably the formal assessments of school students of the UK, the examination authorities are silent on the technical performance of the tests. Examination processes remain arcane with little or no information being made available in relation to reliability or validity, and no recognition of the rights of students to have access to this type of information. Yet, as we mention above, reliability is demonstrably a problem.

The large majority of the 264 standards are characterized in their language by the use of the word *should*. If this was the imperative *should*, and if there was a means by which the standard could be enforced (which there currently is not), then they would underpin legally protected student rights. However, there is little prospect of them being formally enshrined in law and it will likely remain an advisory *should*. The term *rights* is therefore something of a misnomer, and its substitution by a term such as *expectations* is arguably more accurate. Although not immediately obvious, the language of the standards is quite circumspect, carefully written to avoid any charge of being dogmatic or any potential of being used inappropriately, such as might be the case in attempts to override professional judgment. In situations in which assessments are being subjected to legal scrutiny or challenge, the standards are crystal clear about their advisory role: “... the intent of the Standards is to offer guidance for those judgments” (AERA *et al.*, 1999: 4).

The standards may stop short of speaking to legally enforceable rights but they are no less important in informing professional judgment and in their advocacy of fair and good practice (see, e.g., Camara and Lane, 2006). Among the instances of good practice in seeking to protect student interests and rights, are those standards that specifically address issues of information, consent, privacy, and appeals. For example, the standards articulate the view that if any information about an assessment is given to any candidate, it should be given to all of them (S 8.1). The information should cover all important aspects of the assessments including the intended use of the results, scoring criteria, confidentiality, re-sits, and the release of results (S 8.2). The consent of the student to be assessed is seen as a major principle of procedural propriety, though several potential exclusions are offered: when

the testing is mandated by government; when the testing is a routine part of school activities; and when consent is obviously implied. For example, consent may be assumed in an employment-selection context when the act of applying is deemed to include consent to undertake any assessment that is a part of the process (S 8.4). There should be an entitlement to have the privacy of individuals' results maintained (S 8.5) and secure arrangements for storage of the results (S 8.6). In decision situations, the student should be entitled to receive details of the results and scores, and of their interpretation (S 8.9). Such information should also inform students of any recourse to appeals, challenges, and reviews (S 8.10).

While one whole chapter (chapter 8) of the standards is given over to the rights and responsibilities of test-takers, the other chapters also contain standards that speak to students' rights. For example, if the integrity of a test score is challenged, the student should be informed of any rights to appeal and to have legal representation (S 11.11). Standard 13.14 also states that information on the extent of possible misclassification and errors in measurement should be released.

Much of the intention behind developing the standards was to ensure reliable, fair, and valid testing of students and others, particularly in high-stakes settings, but they are nonetheless daunting for the average student, parent, or person who is not *au fait* with assessment terminology. As a means of making the underpinning principles of the standards more accessible, the Joint Committee on Testing Practices (JCTP) issued their 1988 Code of Fair Testing Practices in Education. This was updated in 2004 and "...applies broadly to testing in education (admissions, educational assessment, educational diagnosis, and student placement)" (JCTP, 2004: 2).

In essence, the JCTP code addresses the importance of students (test-takers) having the fullest information possible, including what rights they may be granted in terms of appeals, having their scripts returned, or being able to take re-sits. As with the standards from which they are derived, they are also advisory. However, in combination with the standards they present a formidable array of widely accepted good practices in assessment, to which most courts and legally constituted tribunals will have due regard.

United Nations Convention on the Rights of the Child

As mentioned above, the United Nations Convention on the Rights of the Child is the third axis of possible protection for students' rights in assessment, or at the very least those students who are under 18 years of age, which is the definition of the child under the Convention. At first sight, this might appear an unlikely bastion for

students' rights in assessment, focusing as it does on issues such as children's fundamental rights to life, freedom from abuse, and from recruitment to armies. Ratified by 192 countries, with the remaining two as signatories (the US and Somalia), courts have less leeway to reject rights actions that arise under the relevant articles of the Convention. Aside from the education-specific articles such as 28 and 29, which respectively assert the right to primary education and to the development of personality and talents, the key articles for our purposes are:

1. *Article 3.* In all actions concerning children, whether undertaken by public or private social-welfare institutions, courts of law, administrative authorities, or legislative bodies, the best interests of the child shall be a primary consideration.
2. *Article 12.* States/Parties shall assure to the child who is capable of forming his or her own views the right to express those views freely in all matters affecting the child, the views of the child being given due weight in accordance with the age and maturity of the child.
3. *Article 3.* This has the potential to call into doubt the legality of the testing regimes currently holding sway in the US and England (not, it should be noted, in the three other UK home countries, Scotland, Northern Ireland, and Wales. In these nations, much of the London-led accountability testing has been abandoned as counter-educational by their devolved governments). Many commentators have called the preponderance and nature of the testing into serious question and some researchers are beginning to bring a United Nations Conventions of the Rights of the Child (UNCRC) lens to bear on the issue (e.g., Yates, 2006). Clearly there is an argument that state testing programs in the US, and the national-curriculum tests of England, are not in the best interests of the young students concerned. Probably more to the point, however, is the tenuous validity (and ethical basis) of using such tests of students' learning for a purpose that is some way removed from their individual learning and progression, that is, school and state accountability.

Article 12 brings the matter of assessment propriety into particular focus. High-stakes tests are so called because they have the potential for considerable impact, whether positive or negative, on students' educational or career choices. More often than not, however, the students concerned have little say either in formulating or responding to any aspect of the major decisions that directly affect them and which arise from such tests. The extent to which re-sits are offered, actual annotated test scripts are returned for scrutiny, information on scoring is shared, or feedback on performance is offered, is generally in the gift of the examiners – and not an established right of the students. The AERA Standards and JCTP Code set out clearly what should happen and it is possible that the UNCRC will

provide a legal platform from which to challenge the more arcane of assessment practices today. However, one major obstacle stands in the way – the special place given to academic judgment.

Academic Judgment

Unless there are procedural or technical grounds for challenging a grade or score, it is most likely that students appealing the assessment of their performance in a test will face the dictum, especially in higher education institutions, “We cannot accept challenges to academic judgment.” For example, the UK’s Quality Assurance Agency (with a remit for the quality assurance of teaching and learning in UK universities and colleges) say simply that “[Since] most institutions do not allow appeals against the exercise of academic judgment . . . it would therefore be helpful to define what constitutes an academic judgment” (QAA, 2008: 7). It defines an appeal as “. . . a request for a review of a decision of an academic body charged with decisions on student progression, assessment and awards” (QAA, 2008: 4). Another body, the Office of the Independent Adjudicator for Higher Education (England and Wales) similarly states: “We cannot look at a complaint if . . . it relates to a matter of academic judgment” (OIAHE, 2008) yet it does allow what it terms academic appeals. The adjudicator can recommend compensation and in the examples of the 32 completed cases provided on the website, compensation payments of up to £9000 for students are recorded. Although these are clearly not court-of-law judgments, they are beginning to form a body of case law in the context of university students’ rights in the assessment of their performance.

Disbarring appeals against academic judgment is not a simple expedient on the part of examination authorities, whether university boards of examiners or major testing agencies. Judgments of student performance, other than the purely correct/incorrect dichotomies of aspects of some disciplines such as mathematics, are generally supported by processes such as those that promote inter-judge (inter-rater) reliability. An example of this process is known in the UK as moderation, in which examiners’ grades and scores are checked and calibrated by other examiners. Agreement trials are processes in which a group of examiners take samples of student scripts and grade them together, aiming to develop a connoisseurship among the examiners on the standards to be applied. There are also systems of double marking (where two examiners independently grade the work) and anonymous marking (to ensure that knowledge of the student does not influence the examiner’s grading). The purpose of all such procedural devices is to ensure that the assessments are as reliable as possible and executed fairly across the

test-taking group. Notwithstanding these efforts, however, much of the process remains largely hidden and unknown to the student. Disappointment with their grading will often prompt students to appeal, to be told that only technical or procedural grounds are allowed.

Courts for their part are also notoriously wary of becoming involved in challenging the internal judgments of schools, universities, or examination agencies. Unless there is *prima facie* evidence of malpractice, which is amenable to statutory or common law redress (e.g., some form of discriminatory practice), they will often not proceed. Even if there is an assessment appeal purporting to be a contravention of UNCRC Article 3 or 12, the prospect of a case coming down to the consideration of the reliability or validity of an academic judgment may well jeopardize any potential finding for the student concerned, specifically because it is considered so difficult to mount an effective challenge to academic judgment.

Summary

Students’ rights in relation to the assessment of their learning or performance do not specifically exist in legislation in most jurisdictions. However, the long-standing and highly respected AERA and JCTP guidelines enshrine many expectations for fair treatment. Arguably, these have the intention of prescribing entitlements or rights and the rationale seems eminently reasonable for expectations such as receiving academic information on tests (e.g., the curriculum covered and assessment criteria), receiving technical information on tests (e.g., reliability and validity), having annotated scripts returned, having recourse to appeals, and having opportunities to re-sit a test. Two main factors seem to be in play: a keen sense of natural justice among educators and an increasing academic and professional understanding of the extent to which uncertainty in scores undermines the pursuit of absolute measures of performance.

As time goes on, the hope must be that increased compliance with the AERA Standards and JCTP Code will reduce the potential of unfairness in treatment or outcome in assessment of students. But when needed, there is the increasing prospect of legal strengthening of students’ rights, especially in high-stakes testing. This will most likely occur through bringing to bear national instruments such as the Bill of Rights or international treaties such as the UNCRC, both of which have seen considerable impact in contexts of discrimination in education over recent years.

See also: Assessment and the Evaluation of Institutional Effectiveness; Formative Assessment; Summative Assessment by Teachers.

Bibliography

- AERA, APA, and NCME (1999). *Standards for Educational and Psychological Testing*. American Educational Research Association, American Psychological Association and National Council on Measurement in Education. Washington, DC: American Psychological Association.
- Baker, E. (2007). *The End(s) of Assessment*. Presidential Speech, American Educational Research Association Annual Conference, 11 April, Chicago. <http://www.cmcgc.com/Media/Synch/270409/40/default.htm> (accessed August 2009).
- Black, P. and Wiliam, D. (2006). The reliability of assessments. In Gardner, J. (ed.) *Assessment and Learning*, pp 119–131. Thousand Oaks, CA: Sage.
- Camara, W. J. and Lane, S. (2006). A historical perspective and current views on the standards for educational and psychological testing. *Educational Measurement, Issues and Practice* 25(3), 35–41.
- Durant, D. (2003). A comparative analysis of key stage tests and teacher assessments. *Paper Presented at BERA Annual Conference*, Heriot-Watt University. www.leeds.ac.uk/educol/documents/00003153.htm (accessed August 2009).
- Gardner, J. (ed.) (2006). *Assessment and Learning*. Thousand Oaks, CA: Sage.
- Houston, P. D. (2007). The seven deadly sins of No Child Left Behind. *Phi Delta Kappan* 88(10), 744–749.
- JCTP (2004). *Code of Fair Testing Practices in Education*. Washington, DC: Joint Committee on Testing Practices. <http://www.apa.org/science/FinalCode.pdf> (accessed August 2009).
- Newton, P. and Whetton, C. (2002). Technical and practical implications of the right to appeal test results. National Foundation for Educational Research. Slough: NFER. <http://www.nfer.ac.uk/publications/other-publications/conference-papers/technical-and-practical-implications-of-the-right-to-appeal-test-results.cfm> (accessed August 2009).
- Nichols, S. L., Glass, G. V., and Berliner, D. C. (2005). High stakes testing and student achievement: Problems for the No Child Left Behind Act. Arizona State University: Educational Policy Studies Laboratory. <http://epicpolicy.org/files/EPSSL-0509-105-EPRU.pdf> (accessed August 2009).
- OIAHE (2008). Resolving student complaints, Office of the Independent Adjudicator for Higher Education. <http://www.oiahe.org.uk/summary.asp> (accessed August 2009).
- PWC (2004). *Financial Modelling of the English Exams System 2003–04*. PriceWaterhouseCoopers. http://www.qca.org.uk/qca_9940.aspx (accessed August 2009).
- QAA (2008). Code of practice for the assurance of academic quality and standards in higher education, section 5: Academic appeals and student complaints on academic matters – October 2007. <http://www.qaa.ac.uk/academicinfrastructure/codeofpractice/section5/default.asp> (accessed August 2009).
- Skidmore, P. (2003). *Beyond Measure: Why Educational Assessment is Failing the Test*. London: DEMOS. <http://demos.co.uk/files/beyondmeasure.pdf> (accessed August 2009).
- United Nations Convention on the Rights of the Child (UNCRC) (1990). Office of the High Commissioner for Human Rights. United Nations: New York. <http://www.unhcr.ch/html/menu3/b/k2crc.htm> (accessed August 2009).
- Wallis, C. and Steptoe, S. (2007). How to fix No Child Left Behind. *Time* 169(23), 34.
- Yates, T. L. (2006). A criticism of the No Child Left Behind Act: How the convention on the rights of the child can offer promising reform of education legislation in America, 5. *Whittier Journal of Child and Family Advocacy* 399.

Further Reading

- Cumming, J. J. and Dickson, E. A. (2007). Equity in assessment: Discrimination and disability issues from an Australian legal perspective. *Education and the Law* 19(3/4), 201–220.
- Cumming, J. J. and Mawdsley, R. D. (2005–2006). Student rights and parent rights in education in Australia. *Australia and New Zealand Journal of Law and Education* 10(2);11(1), 37–54.
- Gardner, J. and Cowan, P. (2005). The fallibility of high stakes 11-plus tests in Northern Ireland. *Assessment in Education* 12(2), 145–165.
- Kornhaber, M. L. (2004). Appropriate and inappropriate forms of testing, assessment, and accountability. *Education Policy* 18(1), 45–70.
- Making Good Progress (2007). Department for children, schools and families, London. <http://www.teachernet.gov.uk/mgppilot/> (accessed August 2009).
- Mansell, W. (2007). *Education by Numbers: The Tyranny of Testing*. London: Politico's.
- No Child Left Behind Act 2001 (2002). US Congress Public Law 107–110. <http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf> (accessed August 2009).
- Pitoniak, M. J. and Royer, J. M. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. *Review of Educational Research* 71(1), 53–104.
- Rogosa, D. (2003). How accurate are the STAR national percentile rank scores for individual students? – an interpretive guide: Version 2.0, CAT/6 Survey. <http://www-stat.stanford.edu/~rag/acccguide/> (accessed August 2009).
- The White House (2007). The No Child Left Behind Act: Five years of results for America's children, fact sheet, 8 January 2007. Washington, DC: White House. <http://www.whitehouse.gov/news/releases/2007/01/20070108-6.html> (accessed August 2009).

Relevant Websites

- <http://www.pisa.oecd.org> – Programme for International Student Assessment Paris: OECD.
- <http://timss.bc.edu> – Trends in International Mathematics and Science Study.