

FACULTY OF
EDUCATION &
SOCIAL WORK



THE UNIVERSITY OF
SYDNEY

THE BLIND ASSESSOR: ARE WE CONSTRAINING OR ENRICHING STUDENT LEARNING?

THE BLIND ASSESSOR: ARE WE CONSTRAINING OR ENRICHING STUDENT LEARNING?

MONDAY 22 NOVEMBER 2010

NEW LAW SCHOOL BUILDING, THE UNIVERSITY OF SYDNEY

A SYMPOSIUM PRESENTED AS PART OF THE FACULTY OF EDUCATION AND SOCIAL WORK'S CELEBRATION MARKING 100 YEARS SINCE THE UNIVERSITY OF SYDNEY'S FIRST CHAIR IN EDUCATION WAS APPOINTED.

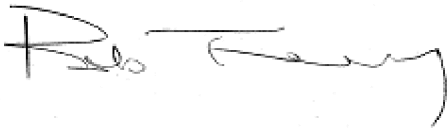
WELCOME TO THE SYMPOSIUM

Welcome to the University of Sydney symposium, *The Blind Assessor: Are we constraining or enriching student learning?* This event is hosted by the Faculty of Education and Social Work and is designed to enrich the debates currently being conducted in the education community and by sections of the broader community about the nature and purpose of systems to monitor, evaluate and improve student learning in our schools.

This year is an important one for the Faculty. In 1910 Alexander Mackie was appointed as the University of Sydney's first professor and chair in education and this symposium is part of the celebrations to mark this important milestone.

Today's program is robust and features some of the leading thinkers and writers in the area of assessment in this country and overseas. We are indeed fortunate to have David Berliner, Regents Professor Emeritus from Arizona State University make the long journey to this country to share with us the considerable body of work that he has done in this area.

Following your participation in the program we will alert you to the availability of papers, PowerPoint handouts and videos on the Faculty website. We thank you for joining us today and hope you find the day stimulating for your own professional context.



Professor Robert J. Tierney
Dean
Faculty of Education and Social Work
The University of Sydney



Margery Hornibrook
Symposium Convener
Faculty of Education and Social Work
The University of Sydney

November 2010

PROGRAM

CHAIR **Peter Freebody**, Faculty of Education and Social Work, The University of Sydney

9.30 WELCOME

LT 101 **Robert J. Tierney**, Dean, Faculty of Education and Social Work, The University of Sydney

9.45 SESSION 1: KEYNOTE

LT 101 STANDARDS, ASSESSMENTS AND THE NARROWING OF THE OUTCOMES OF SCHOOLING AND STUDENTS' MINDS

David C. Berliner, Regents' Professor of Education Emeritus, Arizona State University

10.45 Questions

11.00 MORNING TEA

11.30 SESSION 2: KEYNOTE

LT 101 AN ASSESSMENT OF INFLUENCES, POSITIVE AND NEGATIVE ON EDUCATIONAL ASSESSMENT TODAY

Gabrielle Matters, Principal Research Fellow, Australian Council for Educational Research (ACER)

12.30 Questions

12.45 LUNCH

1.30 SESSION 3: CONCURRENT FEATURES

RM 102 MEASURE TWICE, CUT ONCE

Peter Adams, General Manager Assessment, Australian Curriculum, Assessment and Reporting Authority (ACARA)

LT 101 MANDATED LITERACY ASSESSMENT AND THE REORGANISATION OF TEACHERS' WORK: AN INSTITUTIONAL ETHNOGRAPHY

Barbara Comber, Research Professor, Faculty of Education, Queensland University of Technology

RM 100 STANDARDS SEEM LIKE A GOOD IDEA, BUT HOW DO WE VALIDATE THEM?

Gordon Stanley, Pearson Professor of Educational Assessment, University of Oxford

RM 105 SYMPOSIUM: INDIVIDUAL DIFFERENCES OR INDIVIDUAL DIFFERENCE

Gerald Tindal, Professor and Department Head of Educational Methodology, Policy and Leadership, University of Oregon

2.30 AFTERNOON TEA

2.45 SESSION 4: PANEL DISCUSSION

LT 101 Panel with all six speakers

3.45 CLOSE

SESSION 1

9.45 – 10.45 AM

KEYNOTE

LECTURE THEATRE 101

STANDARDS, ASSESSMENTS AND THE NARROWING OF THE OUTCOMES OF SCHOOLING AND STUDENTS' MINDS

The inevitable responses to high-stakes testing, wherein students' test scores are highly consequential for teachers and administrators, include cheating, excessive test preparation, changes in test scoring, and other forms of gaming to ensure that test scores appear high. But the most pernicious response to high-stakes testing is perhaps the most rational, namely, curriculum narrowing. In this way more of what is believed to be on the test is taught. Curriculum narrowing, however, reduces many students' chances of being thought talented in school and results in a restriction in the creative and enjoyable activities engaged in by teachers and students. The tests commonly used with narrower curricula also appear to restrict thinking skills. In addition, responses to high-stakes environments can easily retard the development of achievement in later grades as a function of the restrictions on learning in earlier grades. Finally, narrowing compromises interpretations of construct validity. The dominance of testing as part of American and British school reform policies insures that many of the skills thought to be most useful in the 21st century will not be taught. Students and their national economies will suffer when nations rely too heavily on high-stakes testing to improve their schools.

DAVID C. BERLINER

ARIZONA STATE UNIVERSITY

David C. Berliner is Regents' Professor of Education Emeritus at Arizona State University, and on the summer faculty of Simon Fraser University in Vancouver, Canada. He has also taught at the Universities of Arizona and Massachusetts, at Teachers College and Stanford University, and at universities in Australia, The Netherlands, Denmark, Spain, and Switzerland. He is a member of the National Academy of Education, the International Academy of Education, and a past president of both the American Educational Research Association (AERA) and the Division of Educational Psychology of the American Psychological Association (APA). He is the winner of numerous awards, most notably the Brock award and the AERA award for distinguished contributions to education, the E. L. Thorndike award from the APA for lifetime achievements, and the NEA 'Friend of Education' award for his work on behalf of the education profession. He is co-author (with B. J. Biddle) of the best seller *The manufactured crisis*, co-author (with Ursula Casanova) of *Putting research to work*, and co-author (with N. L. Gage) of the textbook *Educational psychology*, now in its sixth edition. He is co-editor of the first *Handbook of educational psychology* and the books *Talks to teachers*, and *Perspectives on instructional time*. His newest co-authored book, with Sharon Nichols, is *Collateral damage: How high-stakes testing corrupts American education*. Professor Berliner has authored more than 200 published articles, technical reports, and book chapters.

SESSION 2

11.30 AM – 12.30 PM

KEYNOTE

LECTURE THEATRE 101

AN ASSESSMENT OF INFLUENCES, POSITIVE AND NEGATIVE ON EDUCATIONAL ASSESSMENT TODAY

Joel Stein, writing in *Time* magazine in May this year, listed the 100 least influential people of the year. He made the comment that the task of putting together the least influential list is more difficult than putting together the most influential list. Least influential does not refer to the most negative; it refers to the least influential ideas that used to or ought to have influence. Most influential does not refer to the most positive; it refers to the most influential ideas that ought not to have influence. In the first part of this paper I use these definitions to identify ideas in educational assessment that make up my least influential and most influential lists. Ideas that are discussed and assigned to lists include the following:

- assessment *for* learning; evidence-based policy
- “relevance” in curriculum and assessment
- data and analyses from PISA
- breakthroughs in psychometrics and educational statistics
- advances in test administration and marking
- the paradigm for explaining between-group differences in achievement
- fair use of tests with sub-groups of the test-taking population
- national curriculum and national assessment
- measuring outcomes of schooling
- teachers making a difference
- assessment of higher-order thinking skills
- the validity–reliability trade-off in assessment.

In the second part of this paper I assert that there is little or no evidence that assessment in its present form is assisting more students to learn more of the “big ideas” and of what is intrinsically difficult. And I argue that this state of affairs can be (partially) attributed to the luxury of living in a democracy, the feminization of education, and the length (short) of the political cycle.

GABRIELLE MATTERS

AUSTRALIAN COUNCIL FOR EDUCATIONAL RESEARCH

Gabrielle Matters is a Principal Research Fellow at the Australian Council for Educational Research (ACER) Melbourne, Head of ACER Brisbane, and Executive Secretary of the International Association for Educational Assessment. She is Adjunct Professor, Faculty of Education, Queensland University of Technology, with a doctorate in the field of psychometrics. She holds a Bachelor of Science degree from The University of Queensland and a diploma in piano (AMusA) from the Queensland Conservatorium of Music. Gabrielle is a Fellow of the Australian College of Educators and a recipient of the Queensland Chapter medal. In 2002 she received a Centenary of Federation medal for services to education. Dr Matters has had extensive experience in education as a classroom teacher (physical sciences), school administrator, university lecturer, researcher, advisor, test designer and author. She has held executive management positions within the Australian education sector and has worked with education systems in Australia and overseas, most recently in Tajikistan and Ethiopia. Gabrielle has extensive experience in policy advice and implementation, in working across the academe–government and bureaucracy–school interfaces, and in conceptualising and operationalising initiatives in testing, assessment and examinations. Her areas of expertise are in the fields of educational measurement, educational administration, test design and marking, the curriculum “wars”, the under-achievement of boys, educational research (policy and practice), and system/school reform. She is an author of journal articles, books, and conference papers in these areas and has prepared reports for the Australian government and the Queensland state government.

SESSION 3

1.30 – 2.30 PM

CONCURRENT FEATURE

SEMINAR ROOM 102

MEASURE TWICE, CUT ONCE

The introduction of national testing in literacy and numeracy has proved to be one of the more significant events in educational assessment in Australia, generating widespread discussion and debate. Replacing previous state and territory full cohort testing programs, NAPLAN provides for jurisdiction, school and student level reporting in a national context. Since 2008 over three million students have completed a NAPLAN test series, with more than one million having participated in two series of tests. Considering together the experience of NAPLAN, public commentary on the program, possible future options, and the role of technology, this presentation will explore three questions: *What do we know? What don't we know? What do we need to know?* It also will be put that an 'all or nothing' approach to survey assessments is unhelpful, and that some of the more extreme positions do not recognise the potential for different assessment formats to successfully co-exist and complement one another. A recurring theme will be the importance of measurement and reliable data, and the right type of data, for making informed judgements about student ability and progress. By using a range of assessment formats that target the diversity of student skills and abilities, more informed and better decisions will be made. Drawing upon the building adage "measure twice, cut once", it will be argued that a 'one-off' assessment is not, on its own, a sufficient basis on which to make critical decisions about student ability and progress. However, if that assessment provides high quality data that can be considered together with other assessment information, how could it not be valuable?

PETER ADAMS

AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY

Peter Adams is currently General Manager Assessment at the Australian Curriculum, Assessment and Reporting Authority (ACARA), coming from the position of General Manager, School Programs at the Australian Council for Educational Research (ACER). Prior to ACER Peter project directed ACARA's development of the *My School* website. He also project directed the introduction in 2008 of the National Assessment Program – Literacy and Numeracy (NAPLAN) and the 2009 NAPLAN and equating test series. Peter has held various other roles including: General Manager at Curriculum Corporation; Director Assessment and Testing (Australia) for Pearson *plc*; VCE Examinations Manager at the Victorian Curriculum and Assessment Authority (VCAA); State Reviewer for VCE Art; and Manager of The Royal Australian College of General Practitioners Victoria Faculty. Peter has qualifications in economics, education and postgraduate qualifications in accounting and finance and art education. He also has 18 years experience in teaching and leadership roles within Victorian schools.

SESSION 3

1.30 – 2.30 PM

CONCURRENT FEATURE

LECTURE THEATRE 101

MANDATED LITERACY ASSESSMENT AND THE REORGANISATION OF TEACHERS' WORK: AN INSTITUTIONAL ETHNOGRAPHY

This presentation explores how mandated literacy assessment is reorganising teachers' work in the context of Australia's recently implemented National Assessment Program – Literacy and Numeracy (NAPLAN) which was introduced in 2008. Students in Years 3, 5, 7 and 9 are tested in May and school results are publicly available online in October. The emergence of different forms of educational work associated with the testing phenomenon are identified and discussed. I then report on emergent findings from an institutional ethnography being undertaken in two states of Australia which investigates how the widening insistence on evidence-based teaching is reconfiguring what happens in contrastive school communities. What mandated literacy assessment *does* in a culturally diverse low socioeconomic school community is illustrated in terms of strategic exclusions of students from the testing process, appropriations and adaptations of literacy theory, work intensification, ethical mediation and new rhythms of accountability regimes.

BARBARA COMBER

QUEENSLAND UNIVERSITY OF TECHNOLOGY

Barbara Comber is a Research Professor in the Faculty of Education at the Queensland University of Technology. Her interests include literacy education and social justice, teachers' work and identities, place and space, and practitioner inquiry. She has recently co-edited two books – *The Hawke Legacy: Towards a sustainable society* (Bloustien, Comber & McKinnon, 2009) and *Literacies in place: Teaching environmental communication* (Comber, Nixon & Reid, 2007). She is currently conducting two research projects, one focusing on mandated literacy assessment and the other on new literacy demands in the middle years of schooling.

SESSION 3

1.30 – 2.30 PM

CONCURRENT FEATURE

SEMINAR ROOM 100

STANDARDS SEEM LIKE A GOOD IDEA, BUT HOW DO WE VALIDATE THEM?

Most education systems around the world now report their student performance with reference to standards. This move has been hastened by the impact of international testing and the public policy focus on education in the development of human and social capital. Countries are looking beyond their borders for comparability in student outcomes/results. How to meet this challenge of comparable assessment is an important issue for education systems.

In Europe and in many Commonwealth countries the first approach to comparison has been to develop qualifications frameworks to classify levels of qualifications and to define their common characteristics as outcome standards. Standards describe what it is that students should be able to know and do. While this definition is basically the same across countries and systems; what varies is the “breadth” of the variable (ie. the descriptions); “the name (or type/purpose)”.

Specifying standards on terms of observable student outcomes is not easy. Grading against specific objectives often leads to finer and finer level of specification (‘check-list’ approaches). Danger is that there can be too many to consider and each element may become operationally isolated from each other.

Clearly a fundamental question which arises in education systems is how do we validate standards-referenced results? Public examination authorities are expected to maintain standards. A number of procedures and tools to assist the process in systems where results are referenced to standards will be discussed.

GORDON STANLEY

UNIVERSITY OF OXFORD

Until October this year Gordon was the Inaugural Pearson Professor of Educational Assessment at the University of Oxford and Director, Oxford University Centre for Educational Assessment, and Research Fellow of St Anne's College. Prior to taking up this position at Oxford he spent 10 years heading a curriculum and public examinations authority in New South Wales, Australia. His career has involved teaching and research in assessment as well as holding statutory offices in education. He is an Emeritus Professor of Psychology at the University of Melbourne and Honorary Professor of Education at the University of Sydney. In one way or another his career has involved measuring what many would call the immeasurable. Reliability of public examinations and testing regimes has been of particular interest from both a technical and public interest point of view.

SESSION 3

1.30 – 2.30 PM

CONCURRENT FEATURE

SEMINAR ROOM 105

SYMPOSIUM: INDIVIDUAL DIFFERENCES OR INDIVIDUAL DIFFERENCE

These two terms set the occasion for teachers and administrators to make important decisions and create a tension between large scale testing programs and classroom based assessment systems. The divide is made even wider between these terms when students with disabilities enter the discussion, particularly when the focus is on school accountability and causal explanations. In the symposium, policy and empirical practice are considered in properly aligning the purpose for decision-making and the information being used.

GERALD TINDAL

UNIVERSITY OF OREGON

Dr Tindal is currently the Castle-McIntosh-Knight Professor in the College of Education at the University of Oregon. He is the Department Head of Educational Methodology, Policy, and Leadership and directs Behavioural Research and Teaching (BRT). His research focuses on integrating students with disabilities in general education classrooms using curriculum-based measurement for screening students at risk, monitoring student progress, and evaluating instructional programs. Gerald Tindal also has conducts research on large scale testing and development of alternate assessments. This work includes investigations of teacher decision-making on test participation, test accommodations, and extended assessments of basic skills. He publishes and reviews articles in many special education journals and has written several book chapters and books on curriculum-based measurement and large-scale testing. He teaches courses on assessment systems, data driven decision-making, research design, and program evaluation.

SESSION 4
2.45 – 3.45 PM

PANEL DISCUSSION
LECTURE THEATRE 101

CHAIR

PETER FREEBODY

FACULTY OF EDUCATION AND SOCIAL WORK, THE UNIVERSITY OF SYDNEY
Professor Peter Freebody is a University Professorial Research Fellow based in the Faculty of Education and Social Work. In the past he has worked at the University of New England, Griffith University, and the University of Queensland. Peter also helped to establish the Centre for Research in Pedagogy and Practice at the National Institute of Education in Singapore. He has served on various Australian state and national advisory groups in the area of literacy education, and was Academic Advisor to the Queensland Minister of Education and the Arts. Peter has been a senior consultant to the development of the *Australian Curriculum: English* and to the Curriculum Corporation's national online curriculum program (The Le@rning Federation), and was an advisor to the National Assessment Program in Literacy and Numeracy.

PANEL

DAVID C. BERLINER

REGENTS' PROFESSOR OF EDUCATION EMERITUS, ARIZONA STATE UNIVERSITY

PETER ADAMS

GENERAL MANAGER - ASSESSMENT, AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY (ACARA)

BARBARA COMBER

RESEARCH PROFESSOR, FACULTY OF EDUCATION, QUEENSLAND UNIVERSITY OF TECHNOLOGY

GABRIELLE MATTERS

PRINCIPAL RESEARCH FELLOW, AUSTRALIAN COUNCIL FOR EDUCATIONAL RESEARCH (ACER)

GORDON STANLEY

INAUGURAL PEARSON PROFESSOR OF EDUCATIONAL ASSESSMENT, UNIVERSITY OF OXFORD

GERALD TINDAL

PROFESSOR AND DEPARTMENT HEAD OF EDUCATIONAL METHODOLOGY, POLICY AND LEADERSHIP, UNIVERSITY OF OREGON

PROFESSIONAL LEARNING
Faculty of Education and Social Work
T +61 2 9351 8520
F +61 2 9351 6249
E rachel.payne@sydney.edu.au
sydney.edu.au/education_social_work/professional_learning

FACULTY OF
EDUCATION &
SOCIAL WORK



THE UNIVERSITY OF
SYDNEY

Produced by Professional Learning, Faculty of Education and Social Work,
the University of Sydney, November 2010.
The University reserves the right to make alterations to any information
contained within this publication without notice.

ABN 15 211 513 464
CRICOS 00026A